(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 April 2001 (12.04.2001)

PCT

(10) International Publication Number
WO 01/25431 A1

(51) International Patent Classification[7]: C12N 15/12, C07K 14/705, C12N 5/10, 15/62, C07K 16/28, C12Q 1/68, G01N 33/50

(21) International Application Number: PCT/US00/27211

(22) International Filing Date:
29 September 2000 (29.09.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/157,267	1 October 1999 (01.10.1999)	US
60/225,543	16 August 2000 (16.08.2000)	US
09/640,209	16 August 2000 (16.08.2000)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 09/640,209 (CON)
Filed on 16 August 2000 (16.08.2000)

(71) Applicant *(for all designated States except US)*: THE ROCKEFELLER UNIVERSITY [US/US]; 1230 York Avenue, New York, NY 10021-6399 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: MOMBAERTS, Peter [US/US]; 1230 York Avenue, New York, NY 10021-6399 (US). RODRIGUEZ, Ivan [US/US]; 1230 York Avenue, New York, NY 10021-6399 (US).

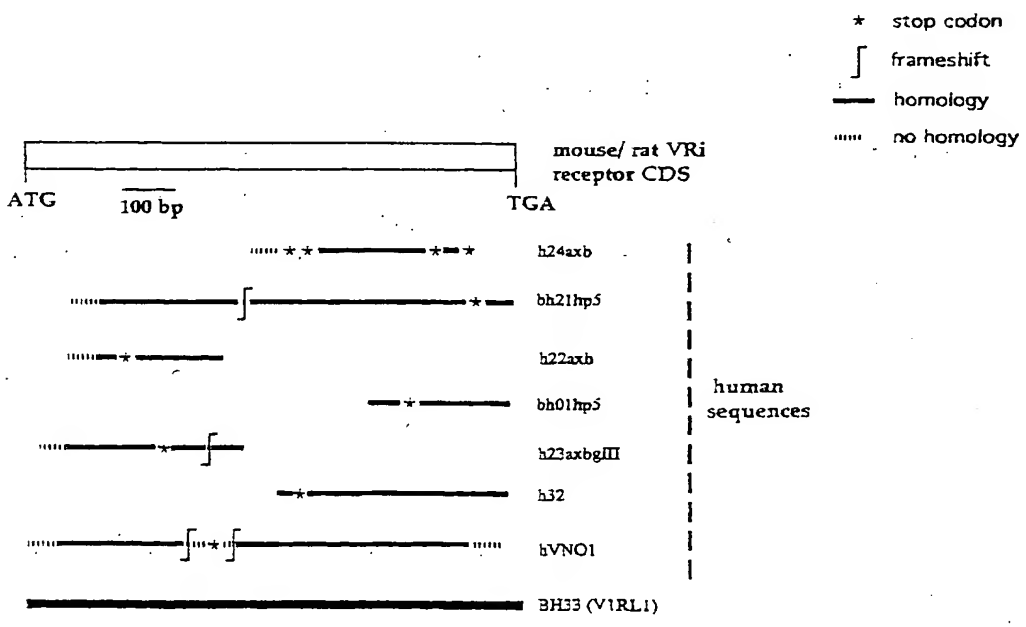
(74) Agents: FEHLNER, Paul, F. et al.; Darby & Darby P.C., 805 Third Avenue, New York, NY 10022-7513 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian

[Continued on next page]

(54) Title: PRIMATE, PARTICULARLY HUMAN, VOMERONASAL-LIKE RECEPTOR



(57) Abstract: Novel human vomeronasal-like receptor, homologous to rat and mouse pheromone receptors, and allelic variants, expression start site variants, and splice variants thereof, are described. Nucleic acids encoding the human vomeronasal-like receptor proteins are also provided. Methods are also provided for producing human vomeronasal-like receptor, detecting expression of these receptors, and screening assays for hVLR1 agonists and antagonists. In particular, three allelic variants of a functional human vomeronasal-like receptor have been identified.

WO 01/25431 A1

patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *With international search report.*

— *Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

PRIMATE, PARTICULARLY HUMAN, VOMERONASAL-LIKE RECEPTOR

FIELD OF THE INVENTION

This invention relates to a novel primate vomeronasal-like receptor, particularly a human vomeronasal-like receptor (hVLR1), which is homologous to putative rat and mouse pheromone receptors, and allelic variants thereof. This invention further relates to nucleic acids encoding the primate vomeronasal-like receptor proteins. The invention also relates to a method for producing primate vomeronasal-like receptor, a method for detecting expression of these receptors, and screening assays for hVLR1 agonists and antagonists.

BACKGROUND OF THE INVENTION

The olfactory system provides sensory information about the chemical composition of the external world. In mammals, olfactory chemoreception initiates at the level of sensory neurons that are located in the main olfactory epithelium (MOE) and the epithelium of the vomeronasal organ (VNO). The MOE mediates the detection of volatile odorants. The VNO mediates mainly the detection of nonvolatile odorants, such as pheromones. These are chemical signals that provide information about gender, dominance, and reproductive status between individuals of the same species (Sorensen, Chem. Sens. 21:245-256, 1996). Pheromones elicit in the recipients innate and stereotyped reproductive and social behaviors, along with profound neuroendocrine and physiological changes.

In mammals, the VNO resides in a blind-ended pouch within the septum of the nose. Axonal projections from the VNO converge to form the vomeronasal nerve and reach target cells within the accessory olfactory bulb. The VNO is exclusively connected to specialized centers of the limbic system, including the vomeronasal amygdala, the bed nucleus of the stria terminalis, and specific nuclei

of the ventromedial hypothalamus involved in reproduction and aggression (Herrada *et al.*, Cell 90:763-773, 1997). Pheromones activate the VNO which results in behavioral and endocrine responses that do not involve higher cognitive centers of the brain (Dulac and Axel, Cell 83:195-206, 1995).

It is now well established that a major proportion (over 70%) of the human odorant receptor repertoire consists of pseudogenes (Rouquier, S. et al. Nat. Genet. 18, 243-250, 1998; Rouquier, S.et al., Proc. Natl. Acad. Sci. USA 97, 2870-2874, 2000; Mombaerts, P. Curr. Opin. Genet. Dev. 9, 315-320, 1999), reflecting perhaps a human's decreased dependence on olfactory cues compared to other mammals. However, the existence of pheromones and a functional vomeronasal system in humans remains controversial (Preti, G. & Wysocki, C.J. Advances in Chemical Signals in Vertebrates (ed. Johnston, R.E.) 315-331 (Plenum Press, New York, 1999)). Despite an undisputed presence of VNO-like structure during early human embryogenesis, it regresses after birth to become vestigial in adults (Humphrey, T. J. Comp. Neurol. 73, 431-468, 1940; Stensaas, L.J. et al., J. Steroid Biochem. Mol. Biol. 39, 553-560, 1991). However, it is inappropriate to consider the VNO as the exclusive site of pheromone detection (Johnston, R.E. Vol.855 (ed. Murphy, C.) 333-348 (Annals of the New York Academy of Sciences, New York, 1998), because some mammals such as the rabbit (Hudson, R. & Distel, H.. Physiol. Behav. 37, 123-128,1986). and the pig (Dorries, K.M., et al., B.P. Brain Behav. Evol. 49, 53-62, 1997) are able to detect pheromones via the main olfactory system; furthermore, fish lack a VNO but express V2R homologs within their olfactory epithelium (Naito, T. et al., Proc. Natl. Acad. Sci. USA 95, 5178-5181, 1998; Cao, Y. et al., Proc. Natl. Acad. Sci. USA 95, 11987-11992, 1998).

The functional and anatomical dichotomy between the main and vomeronasal (or accessory) olfactory systems is further reflected at the level of the molecules that serve as receptors, or putative receptors, for their respective sensory stimuli. In the main olfactory system, odorant receptor genes encode seven-transmembrane proteins and are members of a multigene family that may comprise as many as 1000 genes in rat and mouse (Buck and Axel, Cell 65:175-187, 1991). In the

VNO, two families of genes encoding seven-transmembrane proteins have been proposed to encode pheromone receptors. The first family of vomeronasal receptor genes consists of 30-100 genes that are expressed selectively in vomeronasal sensory neurons of the apical zone of the epithelium of the VNO (Dulac and Axel, *supra*). The second family of vomeronasal receptor genes comprises 30-140 genes that are expressed in vomeronasal sensory neurons of the basal zone (Herrada *et al.*, *supra*). There are no conserved motifs between the two families of vomeronasal receptors, and vomeronasal receptors have no sequence homology with odorant receptors. Neurons in the apical and basal zones express G protein subunits, respectively, $G_{\alpha i2}$ and $G_{\alpha o}$, and project their axons to distinct regions in the mouse accessory olfactory bulb (Berghard and Buck, J. Neurosci. 16:909-918, 1996). The existence of segregated fibers and corresponding G proteins suggests that distinct pheromone signals are likely to elicit electrical stimulation of restricted populations of VNO sensory neurons in order to generate distinct behavioral responses (Herrada *et al.*, *supra*).

It is not known whether G proteins are involved in signal transduction of pheromonal stimuli; they serve as useful molecular markers whose expression tends to correlate with that of the two families of vomeronasal receptor genes. The vomeronasal receptor genes encode putative pheromone receptors; there is no evidence that any of these molecules is a receptor for a pheromone. Because few mammalian pheromones have been identified at the molecular level, ligand-receptor interactions are difficult to define (Sorensen, *supra*).

Human menstrual studies provide strong evidence of the existence of human "pheromones". Studies of women living together in college dormitories report significant increase in synchronization, *i.e.*, a decrease in the difference between onset dates, among roommates and close friends (McClintock, Nature 229:244-245, 1971). In the late follicular phase of women's menstrual cycles, odorless compounds from women's armpits accelerated the preovulatory surge of luteinizing hormone of recipient women and shortened their menstrual cycles. Odorless compounds from the same donors collected later in the menstrual cycle, *i.e.*,

at ovulation, delayed the luteinizing-hormone surge of the recipients and lengthened their menstrual cycles (Stern and McClintock, Nature 392:177-179, 1998). These odorless compounds are postulated as possible candidates for human pheromones.

Recent studies show evidence supportive of a functional vomeronasal system in humans. During prenatal development, the VNO appears distinctly in the submucosa of the septal wall. The fully developed organ has a multistratified epithelium around a narrow lumen (Kjaer and Fischer Hansen, Eur. J. Oral Sci. 104:34-40, 1996). Clinical examinations revealed paired bilateral vomeronasal pits on the anterior third of the nasal septum. The vomeronasal pit leads to a closed tube, lined by a unique pseudostratified columnar epithelium with short microvilli (Morgan *et al.*, J. Steroid Biochem. Molec. Biol. 39:545-552, 1991). Calbindin-like immunoreactivity has been found in epithelial cells of the newborn and adult human vomeronasal organ (Johnson *et al.*, Brain Research 638:329-333, 1994). Some studies suggest that adult human VNO may display species-specific, gender-dimorphic and stereospecific responses to vomeropherin ligands (Monti-Bloch *et al.*, Psychoneuroendocrinology 19(5-7):673-386, 1994).

In U.S. Pat. No. 5,668,006, G-protein linked receptors are reported to control many physiological functions, such as mediating transmembrane signaling from external stimuli (vision, taste and smell), endocrine function (pituitary and adrenal), exocrine function (pancreas), heart rate, lipolysis, and carbohydrate metabolism. The molecular cloning of a number of such receptors have revealed many structural and genetic similarities, permitting classification of the G protein-linked receptor superfamily into five distinct groups.

U.S. Pat. No. 5,691,188, describes how upon binding to the receptor, the receptor presumably undergoes a conformation change leading to activation of the G protein. G proteins are described as being comprised of three subunits: a guanyl-nucleotide binding α subunit; a β subunit; and a γ subunit. G proteins cycle between two forms, depending on whether GDP or GTP is bound to the α subunit. When GDP is bound, the G protein exists as a heterotrimer, the G$\alpha\beta\gamma$ complex. When GTP is bound, the α subunit dissociates, leaving a G$\beta\gamma$ complex. When a G$\alpha\beta\gamma$ complex

operatively associates with an activated G protein coupled receptor in a cell membrane, the rate of exchange of GTP for bound GDP is increased and the rate of dissociation of the bound Gα subunit from the Gαβγ complex increases. The free Gα subunit and Gβγ complex are capable of transmitting a signal to downstream elements of a variety of signal transduction pathways. This fundamental scheme of events forms the basis for a multiplicity of different cell signaling phenomena.

Despite the evidence of a human vomeronasal organ, and the characterization of putative rodent pheromone receptors, to date there is no evidence of a functional human vomeronasal-like receptor. Only human pseudogene homologous to a rat gene encoding a vomeronasal receptor has been reported (Duke and Axel, Cell, 83:195-206, 1995). The present inventors efforts at cloning such a receptor also resulted in identification of numerous pseudogenes. Thus, there is a need in the art to isolate and characterize the structure of a human vomeronasal receptor.

There is a further need to study ligand binding to and activation of such a receptor and to screen for agonists and antagonists.

There is also a need to determine whether such a receptor mediates pheromone receptor activity.

SUMMARY OF THE INVENTION

In one embodiment, the invention provides an isolated human, vomeronasal-like receptor that is homologous to rat and mouse vomeronasal receptors at the amino acid level. The human vomeronasal-like receptor shares homology features found in a large number of functional rodent vomeronasal receptors that are putative pheromone receptors but, unlike earlier identified pseudogenes, this protein is encoded by a gene that is not disrupted. Furthermore, the genomic gene for the hVLR1 lacks introns in the coding region. In specific embodiments, the invention provides allelic variants, splice variants, and alternative start-site variants of the human vomeronasal-like receptor.

In a further embodiment, the invention provides an isolated nucleic

acid comprising a sequence that encodes a functional human, vomeronasal-like receptor. The human embodiment of this nucleic acid is free of 3' and 5' non-coding or non-transcribed genomic sequences. This nucleic acid can be provided as a cDNA or a genome sequence, or alternatively joined to a heterologous nucleic acid, such as an expression vector. Genomic hVLR1 comprises three exons, with the coding region found in the last exon, operably associated with an endogenous promoter.

In yet a further embodiment, the invention provides an isolated chimeric polypeptide comprising an amino acid-sequence of a human vomeronasal-like receptor fused to a heterologous amino acid sequence, such as a signal peptide, an antibody tag, an expression tag, a chromatographic tag, a cytoplasmic signal domain, and a G-protein binding domain.

In yet another embodiment, the invention provides an isolated antigenic fragment of the human vomeronasal-like receptor, as well as an antibody that specifically binds to the receptor.

In another embodiment, the invention provides PCR primers, antisense, ribozyme nucleic acids, and vectors for isolation, cloning, and screening for the hVLR1 receptor.

In a further embodiment, the invention provides a method for isolating, expressing, and screening with the human vomeronasal-like receptor.

In a yet further embodiment, the invention provides for identifying an allelic variant of a gene encoding a human vomeronasal-like receptor. In a specific embodiment, the polymorphism is detected by sequencing.

Still another embodiment of the invention concerns the identification of a primate (chimpanzee) vomeronasal-like receptor gene, the protein encoded thereby, related products, and screens based thereon.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1. Nucleotide and deduced amino acid sequence of "short" human vomeronasal-like receptor, hVLR1 (SEQ ID NO:1 and 2, respectively).

Figures 2a and 2b. Nucleic acid and deduced amino acid sequence of

"long" isoform of the functional vomeronasal-like receptor hVLR1 (SEQ ID NO:3 and 4, respectively).

Figures 3a-3d. Amino acid sequence comparison of human vomeronasal-like receptor hVLR1 with mouse pheromone receptor (emb\CAA73256) (SEQ ID NO:19) and rat VN6 pheromone receptor (pir\I61748) (SEQ ID NO:20).

Figure 4. Comparison of mouse/rat vomeronasal receptor "consensus" and human vomeronasal-like receptor pseudogenes and BH33 (hVLR1).

Figure 5 A. Southern blot of human DNA digested with *EcoRI* (RI) or *HindIII* (H) hybridized with a probe specific for the VLR1 coding sequence. B. The genomic structure of V1RL1.

Figure 6. Schematic of swapping hVLR1 into the mouse VR2 locus.

Figures 7a and 7b. Alignment of the amino acid sequences of two human VRL1 variants (a and b) (SEQ ID NO:15 and 16, respectively) with those of the putative chimpanzee ortholog (cV1RL1) (SEQ ID NO:17) and two mouse V1R sequences, mVR23 (SEQ ID NO:18) and mpr2 (SEQ ID NO:19).

Figure 8. Blot analysis of RT-PCR products hybridized with a VRL1 probe.

DETAILED DESCRIPTION OF THE INVENTION

For the first time, functional cDNA clones encoding a hVLR1 receptor have been isolated and their expression characterized (Figures 1 and 2). Eight different sequences, Bh33, hVNO1, h32, bh21hp5, bho1hp5, h22axh, h23axbglII, h24axb, were found to have strong homologies with the coding regions of mouse or rat vomeronasal receptor genes. One functional gene sequence was isolated based on its homology to the mouse and rat orthologues, particularly at highly conserved positions (consensus domains) (Figure 3). The other seven of these are heavily mutated with multiple frameshifts and stop codons in the coding sequences, indicating that they are pseudogenes in humans (Figure 4). In particular, Bh33 sequence has a complete open reading frame and potentially encodes a protein similar to the mouse or rat vomeronasal receptors from the VR1 family. Many of the conserved amino acids

in rat and mouse vomeronasal receptors are also conserved in Bh33. This discovery is particularly surprising and unexpected since only human pseudogenes have been found to date that bear homology to rat and mouse vomeronasal receptors.

In addition, the present invention provides an N-terminal trunctated splice variant of hVLR1, as depicted in Figure 5B and detected in olfactory epithelium. Southern blot analysis of human genomic DNA with the VLR1 probe reveals a single band (Figure 5A), for both human DNA digested with *EcoRI* (RI) or *HindIII* (H), indicating that the VRL1 gene is not part of a human multigene subfamily of closely related proteins.

A sequence corresponding to Bh33 can be found in GenBank (AC 004076) embedded within a large sequence, namely, a 500 kb ZNF gene family of human chromosome 19, cosmid R30217. The GenBank sequence tentatively identified a partial open reading frame region that is similar to a rat pheromone receptor VN6 (u36898), with 27% identity at the deduced amino acid sequence level. However, the deduced amino acid sequence is not provided, so there is no basis for concluding from this information that the annotated region encodes anything other than a pseudogene.

The term "hVLR1" or "V1RL1" as used herein refers to a primate, preferably human (RVLR1) form of the vomeronasal-like receptor. Such a receptor is characterized by one or more of the following distinct features:

(a) it has two alternative start-site isoforms, a "short" form, which is about a 313 amino acid polypeptide bearing close homology to mouse and rat vomeronasal receptors, and a "long" form, which is about a 353 amino acid polypeptide, both with seven putative transmembrane domains; and an alternative splice form yielding an N-terminal truncated product of about 243 amino acids sharing homology with vomeronasal receptors on the last two-thirds of its amino acid sequence;

(b) it is encoded by a cDNA of about 942 nucleotides ("short" form) or 1059 nucleotides ("long" form);

(c) the amino acid sequence of the short form (and the long form.

over the corresponding region, *i.e.*, excluding the first 40 amino acid residues) is 28% identical and 47% similar to the mouse pheromone receptor (emb\CAA73256) by BLASTX analysis.

In a specific embodiment, the hVLR1 has an amino acid sequence as shown in SEQ ID NO:2 or 4, or any 10 amino acid portion thereof. In another embodiment, the hVLR1 is an allelic variant having the sequence as shown in SEQ ID NO:2, with one or two amino acid differences: Ser201 to Phe; Ala229 to Asp; or both, or any 10 amino acid portion thereof. Corresponding "long" form allelic variants are also contemplated. A VLR1 of the invention will have at least about 90% amino acid sequence identity with SEQ ID NO:2 or 4, preferably about 95% identity, and may have 99% or greater identity, as exemplified by the allelic variants described here. In specific embodiments, chimp VLR1(cVLR1) has 93% amino acid sequence identity with V1RL1a (SEQ ID NO: 15)and V1RL1 b (SEQ ID NO: 16) and human polymorphic variants have 99% identity.

BLAST (Basic Local Alignment Search Tool) program was used to search for homology. Specifically, the search algorithm, BLASTX, was used according to Altschul, et. al., "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 1990, 25:3389-3402, to compare the six-frame conceptual translation products of the hVLR1 942 nucleotide query sequence (both strands) against a protein sequence database. The fundamental unit of BLAST algorithm output is the High-scoring Segment Pair (HSP). An HSP consists of two sequence fragments of arbitrary but equal length whose alignment is locally maximal and for which the alignment score meets or exceeds a threshold or cutoff score. A set of HSPs is thus defined by two sequences, a scoring system, and a cutoff score; this set may be empty if the cutoff score is sufficiently high. In the programmatic implementations of the BLAST, each HSP consists of a segment from the query sequence and one from a database sequence (Altschul *et al.*, *supra*).

The approach to similarity searching taken by the BLAST programs is first to look for similar segments (HSPs) between the query sequence and a database sequence, then to evaluate the statistical significance of any matches that were found.

and finally to report only those matches that satisfy a user-selectable threshold of significance (E parameter). The Expect value (E) is a parameter that describes the number of hits one can "expect" to see just by chance when searching a database of a particular size. It decreases exponentially with the Score (S) that is assigned to a match between two sequences. Essentially, the E value describes the random background noise that exists for matches between sequences. The value is used as a way to create significance threshold for reporting results. When the Expect value is increased from the default value of 10, a larger list with more low-scoring hits can be reported (Altschul *et al.*, *supra*).

Table 1 illustrates the BLASTX results for the hVLR1 shown in SEQ ID NO:2. hVLR1 can be characterized by the relationships shown in this table:

Table 1 - Bh33 BLASTX search results.

Sequence Identifier	Sequence Description	% Identity	% Similarity	Score HSP	E value
emb\CAA73256	mouse pheromone* receptor 2 (Y12724) [Mus musculus]	28	47	104	7e-22
pir\A57223	rat VN3 pheromone receptor [1055248 (U36895)]	28	45	100	2e-20
pir\I61748	rat VN6 pheromone receptor [1055254 (U36898)]	28	44	98	7e-20
pir\I61746	rat VN4 pheromone receptor [1055250 (U36896)]	27	45	96	4e-19
emb\CAA73257	mouse pheromone receptor 1 (Y12725) [Mus musculus]	27	45	90	3e-17
pir\I61749	rat VN2 pheromone receptor [1055256 (U36899)]	27	44	89	6e-17
gi\1039470	VN1 pheromone receptor (U36785) [Rattus norvegicus]	26	43	85	7e-16
gi\1039472	VN7 pheromone receptor (U36786) [Rattus norvegicus]	25	44	83	2e-15
pir\I61747	rat VN5 pheromone receptor [1055252 (U36897)]	28	46	80	2e-14
emb\CAA77091	mouse thyrotropin releasing hormone receptor (Y18244)	29	48	42	0.005

*Although not so designated in the database, all of the receptors in this table are putative pheromone receptor.

hVLR1 can also be characterized by being functional, *i.e.*, it is expressed as a protein. Functionality of hVLR1 of the invention includes binding pheromone-like substrate, and pheromone-like antagonist, which can be affected by GTP; G-protein binding; and signal transduction in response to binding vomeropherin or a vomeropherin agonist. Signal transduction may be evaluated by intracellular calcium mobilization, cyclic AMP accumulation, activation of other G-protein coupled signal transduction pathways, reporter gene expression coupled to G-protein signal transduction, and other methods.

Various chimeric constructs comprising hVLR1 are contemplated as well. Such constructs comprise an hVLR1 fused to a heterologous amino acid sequence, *e.g.*, having functional activity. For example, the hVLR1 can be tagged with an N-terminal or C-terminal tag, such as Myc or FLAG, for immuno-precipitation. Alternatively, a signal sequence can be substituted for the endogenous signal sequence for more efficient processing into the rough endoplasmic reticulum, golgi, and cell membrane. Alternatively, an expression tag, such as an α-mating factor sequence for yeast expression, or residual amino acid residues from a recombinant construct, may be present. In yet another embodiment, a chromatographic tag or handle can be joined to hVLR1. For example, a polyhistidine sequence permits purification on a nickle chelation column. Various combinations of hVLR1 and mouse or rat VNR segments yield alternative chimeric forms of the receptor. Other chimeric constructs, in which heterologous signal transduction domains or G-protein-binding domains are incorporated in the protein are discussed in greater detail, *infra*.

Thus, the present invention advantageously provides a nucleic acid encoding a human vomeronasal-like receptor (hVLR1), the polypeptide encoded by this nucleic acid, cells stably expressing hVLR1, and methods for using such cells, *e.g.*, to screen for hVLR1 agonists and antagonists, particularly agonists and antagonists that are selective for a vomeronasal-like receptor.

In a specific embodiment, the nucleotide sequences encoding the amino acids comprising the novel receptor protein are depicted in SEQ ID NO:1 or

SEQ ID NO:3 for hVLR1, or allelic variants thereof as described above.

Receptors expressed from novel hVLR1 DNAs may be expressed in eukaryotic and prokaryotic cells and can be used to develop and/or implement high throughput screens to identify novel pheromone-like agonists and antagonists. These novel DNAs may be used to help identify receptor subtype selective ligands and may be used to make chimeric and mutant vomeronasal-like receptors which can be used to identify critical ligand binding domains as well as to determine selectivity of ligands. These novel DNAs can be used to further investigate signal transduction systems of vomeronasal-like receptors as well as to determine tissue distribution of receptors.

Translation of hVLR1 cDNAs results in protein sequences which display many of the characteristics of G protein coupled receptors. The peptide sequences of these novel cDNAs may be used to generate antibodies. Antibodies to the receptor can be used to activate the receptor, e.g., by aggregating them.

In a specific embodiment, the present invention provides a vector adapted for expression in a mammalian cell, which comprises the cDNA encoding the functional hVLR1. The term "adapted for expression in a mammalian cell" means that the regulatory elements necessary for the expression of the cDNA in the mammalian cell are present on the plasmid.

The invention further provides a cDNA probe useful for detecting nucleic acid encoding the hVLR1 receptor comprising a nucleic acid molecule of at least about 20 nucleotides having a sequence complementary to a sequence included within the sequence shown in SEQ ID NO:1 or SEQ ID NO:3. It also provides antisense or triple-helix-forming oligonucleotides capable of suppressing expression of hVLR1.

In a specific embodiment, the term "about" or "approximately" means within 20%, preferably within 10%, and more preferably within 5% of a given value or range. Alternatively, and particularly with respect to biological responses, the term "about" means within an order of magnitude, preferably a factor of 5, and more preferably a factor of 2 of a give value.

As used herein, the term "isolated" means that the referenced material is free of components found in the natural environment in which the material is normally found or that the referenced material is present in a heterologous environment. In particular, isolated biological material is free of cellular components. In the case of nucleic acid molecules, an isolated nucleic acid includes a PCR product, an isolated mRNA, a cDNA, or a restriction fragment. In another embodiment, an isolated nucleic acid is preferably excised from the chromosome in which it may be found, and more preferably is no longer joined to non-regulatory, non-coding regions, or to other genes, located upstream or downstream of the gene contained by the isolated nucleic acid molecule when found in the chromosome. In yet another embodiment, the isolated nucleic acid lacks one or more introns. Isolated nucleic acid molecules can be inserted into plasmids, cosmids, artificial chromosomes, and the like. Thus, in a specific embodiment, a cloned or recombinant nucleic acid is an isolated nucleic acid. An isolated protein may be associated with other proteins or nucleic acids, or both, with which it associates in the cell, or with cellular membranes if it is a membrane-associated protein. An isolated organelle, cell, or tissue is removed from the anatomical site in which it is found in an organism. An isolated material may be, but need not be, purified.

The term "purified" as used herein refers to material that has been isolated under conditions that reduce or eliminate unrelated materials, *i.e.*, contaminants. For example, a purified protein is preferably substantially free of other proteins or nucleic acids with which it is associated in a cell; a purified nucleic acid molecule is preferably substantially free of proteins or other unrelated nucleic acid molecules with which it can be found within a cell. As used herein, the term "substantially free" is used operationally, in the context of analytical testing of the material. Preferably, purified material substantially free of contaminants is at least 50% pure; more preferably, at least 90% pure, and more preferably still at least 99% pure. Purity can be evaluated by chromatography, gel electrophoresis, immunoassay, composition analysis, biological assay, and other methods known in the art.

The use of italics indicates a nucleic acid molecule (*e.g.*, hVLR1,

cDNA, gene, etc.); normal text indicates the polypeptide or protein.

In accordance with the present invention there may be employed conventional molecular biology, microbiology, and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, *e.g.*, Sambrook, Fritsch & Maniatis, *Molecular Cloning: A Laboratory Manual*, Second Edition (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (herein "Sambrook *et al.*, 1989"); *DNA Cloning: A Practical Approach*, Volumes I and II (D.N. Glover ed. 1985); *Oligonucleotide Synthesis* (M.J. Gait ed. 1984); *Nucleic Acid Hybridization* [B.D. Hames & S.J. Higgins eds. (1985)]; *Transcription And Translation* [B.D. Hames & S.J. Higgins, eds. (1984)]; *Animal Cell Culture* [R.I. Freshney, ed. (1986)]; *Immobilized Cells And Enzymes* [IRL Press, (1986)]; B.ÊPerbal, *A Practical Guide To Molecular Cloning* (1984); F.M. Ausubel *et al.* (eds.), *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. (1994).

Molecular Biology - Definitions

"Amplification" of DNA as used herein denotes the use of polymerase chain reaction (PCR) to increase the concentration of a particular DNA sequence within a mixture of DNA sequences. For a description of PCR see Saiki *et al.*, Science, 239:487, 1988.

"Chemical sequencing" of DNA denotes methods such as that of Maxam and Gilbert (Maxam-Gilbert sequencing, Maxam and Gilbert, Proc. Natl. Acad. Sci. USA, 74:560, 1977), in which DNA is randomly cleaved using individual base-specific reactions.

"Enzymatic sequencing" of DNA denotes methods such as that of Sanger (Sanger *et al.*, 1977, Proc. Natl. Acad. Sci. USA, 74:5463, 1977), in which a single-stranded DNA is copied and randomly terminated using DNA polymerase, including variations thereof well-known in the art.

As used herein, "sequence-specific oligonucleotides" refers to related sets of oligonucleotides that can be used to detect allelic variations or mutations in the hVLR1 gene.

A "nucleic acid molecule" refers to the phosphate ester polymeric form

of ribonucleosides (adenosine, guanosine, uridine or cytidine; "RNA molecules") or deoxyribonucleosides (deoxyadenosine, deoxyguanosine, deoxythymidine, or deoxycytidine; "DNA molecules"), or any phosphoester analogs thereof, such as phosphorothioates and thioesters, in either single stranded form, or a double-stranded helix. Double stranded DNA-DNA, DNA-RNA and RNA-RNA helices are possible. The term nucleic acid molecule, and in particular DNA or RNA molecule, refers only to the primary and secondary structure of the molecule, and does not limit it to any particular tertiary forms. Thus, this term includes double-stranded DNA found, *inter alia*, in linear (*e.g.*, restriction fragments) or circular DNA molecules, plasmids, and chromosomes. In discussing the structure of particular double-stranded DNA molecules, sequences may be described herein according to the normal convention of giving only the sequence in the 5' to 3' direction along the nontranscribed strand of DNA (*i.e.*, the strand having a sequence homologous to the mRNA). A "recombinant DNA molecule" is a DNA molecule that has undergone a molecular biological manipulation.

A "polynucleotide" or "nucleotide sequence" is a series of nucleotide bases (also called "nucleotides") in DNA and RNA, and means any chain of two or more nucleotides. A nucleotide sequence typically carries genetic information, including the information used by cellular machinery to make proteins and enzymes. These terms include double or single stranded genomic and cDNA, RNA, any synthetic and genetically manipulated polynucleotide, and both sense and anti-sense polynucleotide (although only sense stands are being represented herein). This includes single- and double-stranded molecules, *i.e.*, DNA-DNA, DNA-RNA and RNA-RNA hybrids, as well as "protein nucleic acids" (PNA) formed by conjugating bases to an amino acid backbone. This also includes nucleic acids containing modified bases, for example thio-uracil, thio-guanine and fluoro-uracil.

The polynucleotides herein may be flanked by natural regulatory (expression control) sequences, or may be associated with heterologous sequences, including promoters, internal ribosome entry sites (IRES) and other ribosome binding site sequences, enhancers, response elements, suppressors, signal sequences.

polyadenylation sequences, introns, 5'- and 3'- non-coding regions, and the like. The nucleic acids may also be modified by many means known in the art. Non-limiting examples of such modifications include methylation, "caps", substitution of one or more of the naturally occurring nucleotides with an analog, and internucleotide modifications such as, for example, those with uncharged linkages (*e.g.*, methyl phosphonates, phosphotriesters, phosphoroamidates, carbamates, etc.) and with charged linkages (*e.g.*, phosphorothioates, phosphorodithioates, etc.). Polynucleotides may contain one or more additional covalently linked moieties, such as, for example, proteins (*e.g.*, nucleases, toxins, antibodies, signal peptides, poly-L-lysine, etc.), intercalators (*e.g.*, acridine, psoralen, etc.), chelators (*e.g.*, metals, radioactive metals, iron, oxidative metals, etc.), and alkylators. The polynucleotides may be derivatized by formation of a methyl or ethyl phosphotriester or an alkyl phosphoramidate linkage. Furthermore, the polynucleotides herein may also be modified with a label capable of providing a detectable signal, either directly or indirectly. Exemplary labels include radioisotopes, fluorescent molecules, biotin, and the like.

A "polymorphism" as used herein denotes a variation in the nucleotide sequence of a gene in an individual. Genes that have different nucleotide sequences as a result of a polymorphism are "alleles." A "polymorphic position" is a predetermined nucleotide position within the sequence. In some cases, genetic polymorphisms are reflected by an amino acid sequence variation, and thus a polymorphic position can result in location of a polymorphism in the amino acid sequence at a predetermined position in the sequence of a polypeptide. An individual "homozygous" for a particular polymorphism is one in which both copies of the gene contain the same sequence at the polymorphic position. An individual "heterozygous" for a particular polymorphism is one in which the two copies of the gene contain different sequences at the polymorphic position.

A "polymorphism pattern" as used herein denotes a set of one or more polymorphisms, including without limitation single nucleotide polymorphisms, which may be contained in the sequence of a single gene or a plurality of genes. In the simplest case, a polymorphism pattern can consist of a single nucleotide

polymorphism in only one position of one of two alleles of an individual . However, one has to look at both copies of a gene. A polymorphism pattern that is appropriate for assessing a particular aspect of cardiovascular status (*e.g.*, predisposition to hypertension) need not contain the same number (nor identity, of course) of polymorphisms as a polymorphism pattern that would be appropriate for assessing another aspect of cardiovascular status (*e.g.*, responsivity to ACE inhibitors for control of hypertension). A "test polymorphism pattern" as used herein is a polymorphism pattern determined for a human subject of undefined cardiovascular status. A "reference polymorphism pattern" as used herein is determined from a statistically significant correlation of patterns in a population of individuals with pre-determined cardiovascular status.

The term "host cell" means any cell of any organism that is selected, modified, transformed, grown, or used or manipulated in any way, for the production of a substance by the cell, for example the expression by the cell of a gene, a DNA or RNA sequence, a protein or an enzyme. Host cells can further be used for screening or other assays, as described *infra*.

Proteins and enzymes are made in the host cell using instructions in DNA and RNA, according to the genetic code. Generally, a DNA sequence having instructions for a particular protein or enzyme is "transcribed" into a corresponding sequence of RNA. The RNA sequence in turn is "translated" into the sequence of amino acids which form the protein or enzyme. An "amino acid sequence" is any chain of two or more amino acids. Each amino acid is represented in DNA or RNA by one or more triplets of nucleotides. Each triplet forms a codon, corresponding to an amino acid. For example, the amino acid lysine (Lys) can be coded by the nucleotide triplet or codon AAA or by the codon AAG. (The genetic code has some redundancy, also called degeneracy, meaning that most amino acids have more than one corresponding codon.) Because the nucleotides in DNA and RNA sequences are read in groups of three for protein production, it is important to begin reading the sequence at the correct amino acid, so that the correct triplets are read. The way that a nucleotide sequence is grouped into codons is called the "reading frame."

A "coding sequence" or a sequence "encoding" an expression product, such as a RNA, polypeptide, protein, or enzyme, is a nucleotide sequence that, when expressed, results in the production of that RNA, polypeptide, protein, or enzyme, *i.e.*, the nucleotide sequence encodes an amino acid sequence for that polypeptide, protein or enzyme. A coding sequence for a protein may include a start codon (usually ATG) and a stop codon.

The term "gene", also called a "structural gene" means a DNA sequence that codes for or corresponds to a particular sequence of amino acids which comprise all or part of one or more proteins or enzymes, and may or may not include regulatory DNA sequences, such as promoter sequences, which determine for example the conditions under which the gene is expressed. Some genes, which are not structural genes, may be transcribed from DNA to RNA, but are not translated into an amino acid sequence. Other genes may function as regulators of structural genes or as regulators of DNA transcription.

A "promoter sequence" is a DNA regulatory region capable of binding RNA polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. For purposes of defining the present invention, the promoter sequence is bounded at its 3' terminus by the transcription initiation site and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence will be found a transcription initiation site (conveniently defined for example, by mapping with nuclease S1), as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase.

A coding sequence is "under the control" or "operatively associated with" of transcriptional and translational control sequences in a cell when RNA polymerase transcribes the coding sequence into mRNA, which is then trans-RNA spliced (if it contains introns) and translated into the protein encoded by the coding sequence.

The terms "express" and "expression" mean allowing or causing the information in a gene or DNA sequence to become manifest, for example producing

mRNA or a protein by activating the cellular functions involved in transcription and translation of a corresponding gene or DNA sequence. A DNA sequence is expressed in or by a cell to form an "expression product" such as an mRNA or a protein. The expression product itself, *e.g.* the resulting mRNA or protein, may also be said to be "expressed" by the cell. A protein expression product can be characterized as intracellular, membrane, or secreted. The term "intracellular" means something that is inside a cell. The term "membrane" means something that is in the cell membrane. A substance is "secreted" by a cell if it appears in significant measure outside the cell, from somewhere on or inside the cell.

The term "transfection" means the introduction of a foreign nucleic acid into a cell. The term "transformation" means the introduction of a "foreign" (*i.e.* extrinsic or extracellular) gene, DNA or RNA sequence to a host cell, so that the host cell will express the introduced gene or sequence to produce a desired substance, typically a protein or enzyme coded by the introduced gene or sequence. The introduced gene or sequence may also be called a "cloned" or "foreign" gene or sequence, may include regulatory or control sequences, such as start, stop, promoter, signal, secretion, or other sequences used by a cell's genetic machinery. The gene or sequence may include nonfunctional sequences or sequences with no known function. A host cell that receives and expresses introduced DNA or RNA has been "transformed" and is a "transformant" or a "clone." The DNA or RNA introduced to a host cell can come from any source, including cells of the same genus or species as the host cell, or cells of a different genus or species.

The terms "vector", "cloning vector" and "expression vector" mean the vehicle by which a DNA or RNA sequence (*e.g.* a foreign gene) can be introduced into a host cell, so as to transform the host and promote expression (*e.g.* transcription and translation) of the introduced sequence. Vectors include plasmids, phages, viruses, etc.; they are discussed in greater detail below.

Vectors typically comprise the DNA of a transmissible agent, into which foreign DNA is inserted. A common way to insert one segment of DNA into another segment of DNA involves the use of enzymes called restriction enzymes that

cleave DNA at specific sites (specific groups of nucleotides) called restriction sites. A "cassette" refers to a DNA coding sequence or segment of DNA that codes for an expression product that can be inserted into a vector at defined restriction sites. The cassette restriction sites are designed to ensure insertion of the cassette in the proper reading frame. Generally, foreign DNA is inserted at one or more restriction sites of the vector DNA, and then is carried by the vector into a host cell along with the transmissible vector DNA. A segment or sequence of DNA having inserted or added DNA, such as an expression vector, can also be called a "DNA construct." A common type of vector is a "plasmid", which generally is a self-contained molecule of double-stranded DNA, usually of bacterial origin, that can readily accept additional (foreign) DNA and which can readily introduced into a suitable host cell. A plasmid vector often contains coding DNA and promoter DNA and has one or more restriction sites suitable for inserting foreign DNA. Coding DNA is a DNA sequence that encodes a particular amino acid sequence for a particular protein or enzyme. Promoter DNA is a DNA sequence which initiates, regulates, or otherwise mediates or controls the expression of the coding DNA. Promoter DNA and coding DNA may be from the same gene or from different genes, and may be from the same or different organisms. A large number of vectors, including plasmid and fungal vectors, have been described for replication and/or expression in a variety of eukaryotic and prokaryotic hosts. Non-limiting examples include pKK plasmids (Clonetech), pUC plasmids, pET plasmids (Novagen, Inc., Madison, WI), pRSET or pREP plasmids (Invitrogen, San Diego, CA), or pMAL plasmids (New England Biolabs, Beverly, MA), and many appropriate host cells, using methods disclosed or cited herein or otherwise known to those skilled in the relevant art. Recombinant cloning vectors will often include one or more replication systems for cloning or expression, one or more markers for selection in the host, *e.g.* antibiotic resistance, and one or more expression cassettes.

The term "expression system" means a host cell and compatible vector under suitable conditions, *e.g.* for the expression of a protein coded for by foreign DNA carried by the vector and introduced to the host cell. Common expression

systems include *E. coli* host cells and plasmid vectors, insect host cells and *Baculovirus* vectors, and mammalian host cells and vectors. hVLR1 may be expressed in PC12, COS-1, or C_2C_{12} cells. Other suitable cells include CHO cells, HeLa cells, 293T (human kidney cells), mouse primary myoblasts, and NIH 3T3 cells.

The term "heterologous" refers to a combination of elements not naturally occurring. For example, heterologous DNA refers to DNA not naturally located in the cell, or in a chromosomal site of the cell. Preferably, the heterologous DNA includes a gene foreign to the cell. A heterologous expression regulatory element is a such an element operatively associated with a different gene than the one it is operatively associated with in nature. In the context of the present invention, an hVLR1 gene is heterologous to the vector DNA in which it is inserted for cloning or expression, and it is heterologous to a host cell containing such a vector, in which it is expressed, *e.g.*, a CHO cell.

The terms "mutant" and "mutation" mean any detectable change in genetic material, *e.g.* DNA, or any process, mechanism, or result of such a change. This includes gene mutations, in which the structure (*e.g.* DNA sequence) of a gene is altered, any gene or DNA arising from any mutation process, and any expression product (*e.g.* protein or enzyme) expressed by a modified gene or DNA sequence. The term "variant" may also be used to indicate a modified or altered gene, DNA sequence, enzyme, cell, etc., *i.e.*, any kind of mutant.

"Sequence-conservative variants" of a polynucleotide sequence are those in which a change of one or more nucleotides in a given codon position results in no alteration in the amino acid encoded at that position.

"Function-conservative variants" are those in which a given amino acid residue in a protein or enzyme has been changed without altering the overall conformation and function of the polypeptide, including, but not limited to, replacement of an amino acid with one having similar properties (such as, for example, polarity, hydrogen bonding potential, acidic, basic, hydrophobic, aromatic, and the like). Amino acids with similar properties are well known in the art. For example, arginine, histidine and lysine are hydrophilic-basic amino acids and may be

interchangeable. Similarly, isoleucine, a hydrophobic amino acid, may be replaced with leucine, methionine or valine. Such changes are expected to have little or no effect on the apparent molecular weight or isoelectric point of the protein or polypeptide. Amino acids other than those indicated as conserved may differ in a protein or enzyme so that the percent protein or amino acid sequence similarity between any two proteins of similar function may vary and may be, for example, from 70% to 99% as determined according to an alignment scheme such as by the Cluster Method, wherein similarity is based on the MEGALIGN algorithm. A "function-conservative variant" also includes a polypeptide or enzyme which has at least 60 % amino acid identity as determined by BLAST or FASTA algorithms, preferably at least 75%, most preferably at least 85%, and even more preferably at least 90%, and which has the same or substantially similar properties or functions as the native or parent protein or enzyme to which it is compared.

As used herein, the term "homologous" in all its grammatical forms and spelling variations refers to the relationship between proteins that possess a "common evolutionary origin," including proteins from superfamilies (*e.g.*, the immunoglobulin superfamily) and homologous proteins from different species (*e.g.*, myosin light chain, etc.) (Reeck *et al.*, Cell 50:667, 1987). Such proteins (and their encoding genes) have sequence homology, as reflected by their sequence similarity, whether in terms of percent similarity or the presence of specific residues or motifs at conserved positions.

Accordingly, the term "sequence similarity" in all its grammatical forms refers to the degree of identity or correspondence between nucleic acid or amino acid sequences of proteins that may or may not share a common evolutionary origin (*see* Reeck *et al.*, *supra*). However, in common usage and in the instant application, the term "homologous," when modified with an adverb such as "highly," may refer to sequence similarity and may or may not relate to a common evolutionary origin.

In a specific embodiment, two DNA sequences are "substantially homologous" or "substantially similar" when at least about 80%, and most preferably

at least about 90 or 95%, of the nucleotides match over the defined length of the DNA sequences, as determined by sequence comparison algorithms, such as BLAST, FASTA, DNA Strider, etc. An example of such a sequence is an allelic variant of the specific hVLR1 genes of the invention. Sequences that are substantially homologous can be identified by comparing the sequences using standard software available in sequence data banks, or in a Southern hybridization experiment under, for example, stringent conditions as defined for that particular system.

Similarly, in a particular embodiment, two amino acid sequences are "substantially homologous" or "substantially similar" when greater than 80% of the amino acids are identical, or greater than about 90% are similar (functionally identical). Preferably, the similar or homologous sequences are identified by alignment using, for example, the GCG (Genetics Computer Group, Program Manual for the GCG Package, *Version 7*, Madison, Wisconsin) pileup program, or any of the programs described above (BLAST, FASTA).

A nucleic acid molecule is "hybridizable" to another nucleic acid molecule, such as a cDNA, genomic DNA, or RNA, when a single stranded form of the nucleic acid molecule can anneal to the other nucleic acid molecule under the appropriate conditions of temperature and solution ionic strength (*see* Sambrook *et al.*, *supra*). The conditions of temperature and ionic strength determine the "stringency" of the hybridization. For preliminary screening for homologous nucleic acids, low stringency hybridization conditions, corresponding to a T_m (melting temperature) of 55°C, can be used, *e.g.*, 5x SSC, 0.1% SDS, 0.25% milk, and no formamide; or 30% formamide, 5x SSC, 0.5% SDS). Moderate stringency hybridization conditions correspond to a higher T_m, *e.g.*, 40% formamide, with 5x or 6x SCC. High stringency hybridization conditions correspond to the highest T_m, *e.g.*, 50% formamide, 5x or 6x SCC. SCC is a 0.15M NaCl, 0.015M Na-citrate. Hybridization requires that the two nucleic acids contain complementary sequences, although depending on the stringency of the hybridization, mismatches between bases are possible. The appropriate stringency for hybridizing nucleic acids depends on the length of the nucleic acids and the degree of complementation, variables well known

in the art. The greater the degree of similarity or homology between two nucleotide sequences, the greater the value of T_m for hybrids of nucleic acids having those sequences. The relative stability (corresponding to higher T_m) of nucleic acid hybridizations decreases in the following order: RNA:RNA, DNA:RNA, DNA:DNA. For hybrids of greater than 100 nucleotides in length, equations for calculating T_m have been derived (*see* Sambrook *et al.*, *supra*, 9.50-9.51). For hybridization with shorter nucleic acids, *i.e.*, oligonucleotides, the position of mismatches becomes more important, and the length of the oligonucleotide determines its specificity (*see* Sambrook *et al.*, *supra*, 11.7-11.8). A minimum length for a hybridizable nucleic acid is at least about 10 nucleotides; preferably at least about 15 nucleotides; and more preferably the length is at least about 20 nucleotides.

In a specific embodiment, the term "standard hybridization conditions" refers to a T_m of 55°C, and utilizes conditions as set forth above. In a preferred embodiment, the T_m is 60°C; in a more preferred embodiment, the T_m is 65°C. In a specific embodiment, "high stringency" refers to hybridization and/or washing conditions at 68°C in 0.2XSSC, at 42°C in 50% formamide, 4XSSC, or under conditions that afford levels of hybridization equivalent to those observed under either of these two conditions.

As used herein, the term "oligonucleotide" refers to a nucleic acid, generally of at least 10, preferably at least 15, and more preferably at least 20 nucleotides, preferably no more than 100 nucleotides, that is hybridizable to a genomic DNA molecule, a cDNA molecule, or an mRNA molecule encoding a gene, mRNA, cDNA, or other nucleic acid of interest. Oligonucleotides can be labeled, *e.g.*, with ^{32}P-nucleotides or nucleotides to which a label, such as biotin, has been covalently conjugated. In one embodiment, a labeled oligonucleotide can be used as a probe to detect the presence of a nucleic acid. In another embodiment, oligonucleotides (one or both of which may be labeled) can be used as PCR primers, either for cloning full length or a fragment of hVLR1, or to detect the presence of nucleic acids encoding hVLR1. In a further embodiment, an oligonucleotide of the invention can form a triple helix with a hVLR1 DNA molecule. Generally,

oligonucleotides are prepared synthetically, preferably on a nucleic acid synthesizer. Accordingly, oligonucleotides can be prepared with non-naturally occurring phosphoester analog bonds, such as thioester bonds, etc.

The present invention provides antisense nucleic acids (including ribozymes), which may be used to inhibit expression of hVLR1 of the invention. An "antisense nucleic acid" is a single stranded nucleic acid molecule which, on hybridizing under cytoplasmic conditions with complementary bases in an RNA or DNA molecule, inhibits the latter's role. If the RNA is a messenger RNA transcript, the antisense nucleic acid is a countertranscript or mRNA-interfering complementary nucleic acid. As presently used, "antisense" broadly includes RNA-RNA interactions, RNA-DNA interactions, ribozymes and RNase-H mediated arrest. Antisense nucleic acid molecules can be encoded by a recombinant gene for expression in a cell (*e.g.*, U.S. Patent No. 5,814,500; U.S. Patent No. 5,811,234), or alternatively they can be prepared synthetically (*e.g.*, U.S. Patent No. 5,780,607).

Specific non-limiting examples of synthetic oligonucleotides envisioned for this invention include oligonucleotides that contain phosphorothioates, phosphotriesters, methyl phosphonates, short chain alkyl, or cycloalkl intersugar linkages or short chain heteroatomic or heterocyclic intersugar linkages. Most preferred are those with CH_2-NH-O-CH_2, CH_2-N(CH_3)-O-CH_2, CH_2-O-N(CH_3)-CH_2, CH_2-N(CH_3)-N(CH_3)-CH_2 and O-N(CH_3)-CH_2-CH_2 backbones (where phosphodiester is O-PO_2-O-CH_2). US Patent No. 5,677,437 describes heteroaromatic olignucleoside linkages. Nitrogen linkers or groups containing nitrogen can also be used to prepare oligonucleotide mimics (U.S. Patents No. 5,792,844 and No. 5,783,682). US Patent No. 5,637,684 describes phosphoramidate and phosphorothioamidate oligomeric compounds. Also envisioned are oligonucleotides having morpholino backbone structures (U.S. Pat. No. 5,034,506). In other embodiments, such as the peptide-nucleic acid (PNA) backbone, the phosphodiester backbone of the oligonucleotide may be replaced with a polyamide backbone, the bases being bound directly or indirectly to the aza nitrogen atoms of the polyamide backbone (Nielsen *et al.*, Science 254:1497, 1991). Other synthetic oligonucleotides

may contain substituted sugar moieties comprising one of the following at the 2' position: OH, SH, SCH_3, F, OCN, $O(CH_2)_nNH_2$ or $O(CH_2)_nCH_3$ where n is from 1 to about 10; C_1 to C_{10} lower alkyl, substituted lower alkyl, alkaryl or aralkyl; Cl; Br; CN; CF_3; OCF_3; O-; S-, or N-alkyl; O-, S-, or N-alkenyl; $SOCH_3$; SO_2CH_3; ONO_2; NO_2; N_3; NH_2; heterocycloalkyl; heterocycloalkaryl; aminoalkylamino; polyalkylamino; substituted silyl; a fluorescein moiety; an RNA cleaving group; a reporter group; an intercalator; a group for improving the pharmacokinetic properties of an oligonucleotide; or a group for improving the pharmacodynamic properties of an oligonucleotide, and other substituents having similar properties. Oligonucleotides may also have sugar mimetics such as cyclobutyls or other carbocyclics in place of the pentofuranosyl group. Nucleotide units having nucleosides other than adenosine, ytidine, guanosine, thymidine and uridine, such as inosine, may be used in an oligonucleotide molecule.

Nucleic Acids Encoding hVLR1 Proteins

The present invention contemplates isolation of a gene encoding a hVLR1 of the invention, including a full length, or naturally occurring form of hVLR1, allelic variants and splice variants thereof, and any antigenic fragments thereof from any human source.

A gene encoding hVLR1, whether genomic DNA or cDNA, can be isolated from any source, particularly from a human cDNA or genomic library. Methods for obtaining hVLR1 gene are well known in the art, as described above (*see, e.g.*, Sambrook *et al.*, 1989, *supra*). The DNA may be obtained by standard procedures known in the art from cloned DNA (*e.g.*, a DNA "library"), and preferably is obtained from a cDNA library prepared from tissues with high level expression of the protein (*e.g.*, an olfactory epithelium library, since these are the cells that evidence high levels of expression of hVLR1), by chemical synthesis, by cDNA cloning, or by the cloning of genomic DNA, or fragments thereof, purified from the desired cell (See, for example, Sambrook *et al.*, 1989, *supra*; Glover, D.M. (ed.), 1985, DNA Cloning: A Practical Approach, MRL Press, Ltd., Oxford, U.K. Vol. I, II). Clones derived from genomic DNA may contain regulatory DNA regions in addition to

coding regions. Whatever the source, the gene may be molecularly cloned into a suitable vector for propagation of the gene. Identification of the specific DNA fragment containing the desired hVLR1 gene may be accomplished in a number of ways. For example, a portion of a hVLR1 gene exemplified *infra* can be purified and labeled to prepare a labeled probe, and the generated DNA may be screened by nucleic acid hybridization to the labeled probe (Benton and Davis, Science 196:180, 1977; Grunstein and Hogness, Proc. Natl. Acad. Sci. U.S.A. 72:3961, 1975). Those DNA fragments with substantial homology to the probe, such as an allelic variant from another individual, will hybridize. In a specific embodiment, highest stringency hybridization conditions are used to identify a homologous hVLR1 gene.

Further selection can be carried out on the basis of the properties of the gene, *e.g.*, if the gene encodes a protein product having the isoelectric, electrophoretic, amino acid composition, partial or complete amino acid sequence, antibody binding activity, or ligand binding profile of hVLR1 protein as disclosed herein. Thus, the presence of the gene may be detected by assays based on the physical, chemical, immunological, or functional properties of its expressed product.

Identification of Polymorphisms

The invention specifically contemplates isolating and characterizing allelic variants of hVLR1 having various polymorphisms, and splice variants.

An individual's polymorphisms pattern can be established, *e.g.*, by obtaining DNA from the individual and determining the sequence at a predetermined polymorphic position or positions in a gene, or more than one gene.

The DNA may be obtained from any cell source. Non-limiting examples of cell sources available in clinical practice include without limitation blood cells, buccal cells, cervicovaginal cells, epithelial cells from urine, fetal cells, or any cells present in tissue obtained by biopsy. Cells may also be obtained from body fluids, including without limitation blood, saliva, sweat, urine, cerebrospinal fluid, feces, and tissue exudates at the site of infection or inflammation. DNA is extracted from the cell source or body fluid using any of the numerous methods that are standard in the art. It will be understood that the particular method used to extract

DNA will depend on the nature of the source.

Determination of the sequence of the extracted DNA at polymorphic positions is achieved by any means known in the art, including but not limited to direct sequencing, hybridization with allele-specific oligonucleotides, allele-specific PCR, ligase-PCR, HOT cleavage, denaturing gradient gel electrophoresis (DGGE), and single-stranded conformational polymorphism (SSCP). Direct sequencing may be accomplished by any method, including without limitation chemical sequencing, using the Maxam-Gilbert method; by enzymatic sequencing, using the Sanger method; mass spectrometry sequencing; and sequencing using a chip-based technology. See, *e.g.*, Little *et al.*, Genet. Anal. 6:151, 1996. Preferably, DNA from a subject is first subjected to amplification by polymerase chain reaction (PCR) using specific amplification primers.

In an alternate embodiment, biopsy tissue is obtained from a subject. Antibodies that are capable of distinguishing between different polymorphic forms of hVLR1 are then applied to samples of the tissue to determine the presence or absence of a polymorphic form specified by the antibody. The antibodies may be polyclonal or monoclonal, preferably monoclonal. Measurement of specific antibody binding to cells may be accomplished by any known method, *e.g.*, quantitative flow cytometry, or enzyme-linked or fluorescence-linked immunoassay. The presence or absence of a particular polymorphism and its allelic distribution (*i.e.*, homozygosity vs. heterozygosity) is determined by comparing the values obtained from a patient with norms established from populations of patients having known polymorphic patterns.

In another alternate embodiment, RNA is isolated from biopsy tissue using standard methods well known to those of ordinary skill in the art, such as guanidium thiocyanate-phenol-chloroform extraction (Chomocyznski *et al.*, 1987, Anal. Biochem., 162:156.) The isolated RNA is then subjected to coupled reverse transcription and amplification by polymerase chain reaction (RT-PCR), using specific oligonucleotide primers that are specific for a selected polymorphism. Conditions for primer annealing are chosen to ensure specific reverse transcription and amplification; thus, the appearance of an amplification product is diagnostic of

the presence of a particular polymorphism. In another embodiment, RNA is reverse-transcribed and amplified, after which the amplified sequences are identified by, *e.g.*, direct sequencing. In still another embodiment, cDNA obtained from the RNA can be cloned and sequenced to identify a polymorphism.

hVLR1 Analogs

The present invention also relates to cloning vectors containing genes encoding analogs and derivatives of hVLR1 of the invention, that have the same or homologous functional activity as hVLR1. The production and use of derivatives and analogs related to hVLR1 are within the scope of the present invention. In a specific embodiment, the derivative or analog is functionally active, *i.e.*, capable of exhibiting one or more functional activities associated with a full-length, wild-type hVLR1 of the invention. Such functions include pheromone or other ligand binding, G protein binding and activation, and localization to the cell membrane. In another embodiment, an hVLR1 chimeric construct containing a different cytoplasmic domain, *e.g.*, having an intracellular signaling sequence from another receptor protein, can be prepared. Other chimeric or fusion proteins are also contemplated. Examples include chimeric proteins with G-protein binding domains from the other G protein coupled receptors (GPCRs), GFP fusions, epitope tagged proteins, etc.

hVLR1 derivatives can be made by altering encoding nucleic acid sequences by substitutions, additions or deletions that provide for functionally equivalent molecules. In a specific embodiment, *infra*, a deletion derivative of hVLR1 is prepared and found to have ligand binding and signal transduction properties in the assays used to evaluate the proteins. Preferably, derivatives are made that have enhanced or increased functional activity relative to native hVLR1. Alternatively, such derivatives may encode soluble fragments of hVLR1, or fragments of hVLR1 that contain the extracellular domain that have the same or greater affinity for pheromone-like substrates or other ligands of hVLR1.

Due to the degeneracy of nucleotide coding sequences, other DNA sequences that encode substantially the same amino acid sequence as a hVLR1 gene may be used in the practice of the present invention. These include but are not limited

to allelic genes and nucleotide sequences comprising all or portions of hVLR1 genes which are altered by the substitution of different codons that encode the same amino acid residue within the sequence, thus producing a silent change (sequence conservative variants). Likewise, the hVLR1 derivatives of the invention include, but are not limited to, those containing, as a primary amino acid sequence, all or part of the amino acid sequence of a hVLR1 protein including altered sequences in which functionally equivalent amino acid residues are substituted for residues within the sequence resulting in a conservative amino acid substitution (*e.g.*, functional conservative variants). For example, one or more amino acid residues within the sequence can be substituted by another amino acid of a similar polarity and, if present, charge, which acts as a functional equivalent, resulting in a silent alteration. Substitutes for an amino acid within the sequence may be selected from other members of the class to which the amino acid belongs. For example, the nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan and methionine. Amino acids containing aromatic ring structures are phenylalanine, tryptophan, and tyrosine. The polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine. The positively charged (basic) amino acids include arginine, lysine and histidine. The negatively charged (acidic) amino acids include aspartic acid and glutamic acid. Such alterations will not be expected to affect apparent molecular weight as determined by polyacrylamide gel electrophoresis, or isoelectric point. Particularly preferred substitutions are:

- Lys for Arg and vice versa such that a positive charge may be maintained;
- Glu for Asp and vice versa such that a negative charge may be maintained;
- Ser for Thr such that a free -OH can be maintained; and
- Gln for Asn such that a free $CONH_2$ can be maintained.

Amino acid substitutions may also be introduced to substitute an amino acid with a particularly preferable property. For example, a Cys may be introduced a potential site for disulfide bridges with another Cys.

The genes encoding hVLR1 derivatives and analogs of the invention

can be produced by various methods known in the art. The manipulations which result in their production can occur at the gene or protein level. For example, the cloned hVLR1 gene sequence can be modified by any of numerous strategies known in the art (Sambrook *et al.*, 1989, *supra*). The sequence can be cleaved at appropriate sites with restriction endonuclease(s), followed by further enzymatic modification if desired, isolated, and ligated *in vitro*. In the production of the gene encoding a derivative or analog of hVLR1, care should be taken to ensure that the modified gene remains within the same translational reading frame as the hVLR1 gene, uninterrupted by translational stop signals, in the gene region where the desired activity is encoded.

Additionally, the hVLR1-encoding nucleic acid sequence can be mutated *in vitro* or *in vivo*, to create and/or destroy translation, initiation, and/or termination sequences, or to create variations in coding regions and/or form new restriction endonuclease sites or destroy preexisting ones, to facilitate further *in vitro* modification. In the Examples, *infra*, such modifications were made to introduce restriction sites and facilitate cloning the hVLR1 gene into an expression vector. Any technique for mutagenesis known in the art can be used, including but not limited to, *in vitro* site-directed mutagenesis (Hutchinson, C., *et al.*, J. Biol. Chem. 253:6551, 1978; Zoller and Smith, DNA 3:479-488, 1984; Oliphant *et al.*, Gene 44:177, 1986; Hutchinson *et al.*, Proc. Natl. Acad. Sci. U.S.A. 83:710, 1986), use of TAB™ linkers (Pharmacia), etc. PCR techniques are preferred for site directed mutagenesis (see Higuchi, 1989, "Using PCR to Engineer DNA", in *PCR Technology: Principles and Applications for DNA Amplification*, H. Erlich, ed., Stockton Press, Chapter 6, pp. 61-70).

The identified and isolated gene can then be inserted into an appropriate cloning vector. A large number of vector-host systems known in the art may be used. Possible vectors include, but are not limited to, plasmids or modified viruses, but the vector system must be compatible with the host cell used. Examples of vectors include, but are not limited to, *E. coli*, bacteriophages such as lambda derivatives, or plasmids such as pBR322 derivatives or pUC plasmid derivatives, *e.g.*, pGEX vectors, pmal-c, pFLAG, etc. The insertion into a cloning vector can, for

example, be accomplished by ligating the DNA fragment into a cloning vector which has complementary cohesive termini. However, if the complementary restriction sites used to fragment the DNA are not present in the cloning vector, the ends of the DNA molecules may be enzymatically modified. Alternatively, any site desired may be produced by ligating nucleotide sequences (linkers) onto the DNA termini; these ligated linkers may comprise specific chemically synthesized oligonucleotides encoding restriction endonuclease recognition sequences.

Recombinant molecules can be introduced into host cells via transformation, transfection, infection, electroporation, etc., so that many copies of the gene sequence are generated. Preferably, the cloned gene is contained on a shuttle vector plasmid, which provides for expansion in a cloning cell, *e.g.*, *E. coli*, and facile purification for subsequent insertion into an appropriate expression cell line, if such is desired. For example, a shuttle vector, which is a vector that can replicate in more than one type of organism, can be prepared for replication in both *E. coli* and *Saccharomyces cerevisiae* by linking sequences from an *E. coli* plasmid with sequences from the yeast 2μ plasmid.

Expression of hVLR1

The nucleotide sequence coding for hVLR1, or antigenic fragment, derivative or analog thereof, or a functionally active derivative, including a chimeric protein, thereof, can be inserted into an appropriate expression vector, *i.e.*, a vector which contains the necessary elements for the transcription and translation of the inserted protein-coding sequence. Thus, the nucleic acid encoding hVLR1 of the invention is operationally associated with a promoter in an expression vector of the invention. Both cDNA and genomic sequences can be cloned and expressed under control of such regulatory sequences. An expression vector also preferably includes a replication origin.

Alternatively, an hVLR1 polypeptide of the invention can be prepared using well-known techniques in peptide synthesis, including solid phase synthesis (using, *e.g.*, BOC of FMOC chemistry), or peptide condensation techniques.

As used herein, the terms "polypeptide" and "protein" may be used

interchangably to refer to the gene product (or corresponding synthetic product) of an hVLR1 gene. The term "protein" may also refer specifically to the polypeptide as expressed in cells. A peptide is generally a fragment of a polypeptide, *e.g.*, of about six or more amino acid residues.

The necessary transcriptional and translational signals can be provided on a recombinant expression vector, or they may be supplied by the native gene encoding hVLR1 and/or its flanking regions.

Potential host-vector systems include but are not limited to mammalian cell systems infected with virus (*e.g.*, vaccinia virus, adenovirus, adeno-associated virus, herpes virus, etc.); insect cell systems infected with virus (*e.g.*, baculovirus); microorganisms such as yeast containing yeast vectors; or bacteria transformed with bacteriophage, DNA, plasmid DNA, or cosmid DNA. The expression elements of vectors vary in their strengths and specificities. Depending on the host-vector system utilized, any one of a number of suitable transcription and translation elements may be used.

A preferred expression host is a eukaryotic cell (*e.g.*, yeast, insect, or mammalian cell). More preferred is a mammalian cell, *e.g.*, human, rat, monkey, dog, or hamster cell. In specific embodiments, *infra*, hVLR1 is expressed in a human neuroblastoma cell line (*e.g.*, SK-N-MC), or a chinese hamster ovary cell line (*e.g.*, CHO 293).

A recombinant hVLR1 protein of the invention, or functional fragment, derivative, chimeric construct, or analog thereof, may be expressed chromosomally, after integration of the coding sequence by recombination. In this regard, any of a number of amplification systems may be used to achieve high levels of stable gene expression (*See* Sambrook *et al.*, 1989, *supra*).

Any of the methods previously described for the insertion of DNA fragments into a cloning vector may be used to construct expression vectors containing a gene consisting of appropriate transcriptional/translational control signals and the protein coding sequences. These methods may include *in vitro* recombinant DNA and synthetic techniques and *in vivo* recombination (genetic recombination).

Expression of hVLR1 protein may be controlled by any promoter/enhancer element known in the art, but these regulatory elements must be functional in the host selected for expression. Promoters which may be used to control hVLR1 gene expression include, but are not limited to, cytomegalovirus (CMV) promoter (U.S. Patent Nos. 5,385,839 and 5,168,062), the SV40 early promoter region (Benoist and Chambon, 1981, Nature 290:304-310), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto, *et al.*, Cell 22:787-797, 1980), the herpes thymidine kinase promoter (Wagner *et al.*, Proc. Natl. Acad. Sci. U.S.A. 78:1441-1445, 1981), the regulatory sequences of the metallothionein gene (Brinster *et al.*, Nature 296:39-42, 1982); prokaryotic expression vectors such as the β-lactamase promoter (Villa-Kamaroff, *et al.*, Proc. Natl. Acad. Sci. U.S.A. 75:3727-3731, 1978), or the *tac* promoter (DeBoer, *et al.*, Proc. Natl. Acad. Sci. U.S.A. 80:21-25, 1983); see also "Useful proteins from recombinant bacteria" in Scientific American, 242:74-94, 1980; promoter elements from yeast or other fungi such as the Gal 4 promoter, the ADC (alcohol dehydrogenase) promoter, PGK (phosphoglycerol kinase) promoter, alkaline phosphatase promoter; and the animal transcriptional control regions, which exhibit tissue specificity and have been utilized in transgenic animals: elastase I gene control region which is active in pancreatic acinar cells (Swift *et al.*, Cell 38:639-646, 1984; Ornitz *et al.*, Cold Spring Harbor Symp. Quant. Biol. 50:399-409, 1986; MacDonald, Hepatology 7:425-515, 1987); insulin gene control region which is active in pancreatic beta cells (Hanahan, Nature 315:115-122, 1985), immunoglobulin gene control region which is active in lymphoid cells (Grosschedl *et al.*, Cell 38:647-658, 1984; Adames *et al.*, Nature 318:533-538, 1985; Alexander *et al.*, Mol. Cell. Biol. 7:1436-1444, 1987), mouse mammary tumor virus control region which is active in testicular, breast, lymphoid and mast cells (Leder *et al.*, Cell 45:485-495, 1986), albumin gene control region which is active in liver (Pinkert *et al.*, Genes and Devel. 1:268-276, 1987), alpha-fetoprotein gene control region which is active in liver (Krumlauf *et al.*, Mol. Cell. Biol. 5:1639-1648, 1985; Hammer *et al.*, Science 235:53-58, 1987), alpha 1-antitrypsin gene control region which is active in the liver (Kelsey *et al.*, Genes and

Devel. 1:161-171, 1987), beta-globin gene control region which is active in myeloid cells (Mogram *et al.*, Nature 315:338-340, 1985; Kollias *et al.*, Cell 46:89-94, 1986), myelin basic protein gene control region which is active in oligodendrocyte cells in the brain (Readhead *et al.*, Cell 48:703-712, 1987), myosin light chain-2 gene control region which is active in skeletal muscle (Sani, Nature 314:283-286, 1985), and gonadotropic releasing hormone gene control region which is active in the hypothalamus (Mason *et al.*, Science 234:1372-1378, 1986).

Expression Vectors

A wide variety of host/expression vector combinations (*i.e.*, expression systems) may be employed in expressing the DNA sequences of this invention. Useful expression vectors, for example, may consist of segments of chromosomal, non-chromosomal and synthetic DNA sequences. Suitable vectors include derivatives of SV40 and known bacterial plasmids, e.g., *E. coli* plasmids col El, pCR1, pBR322, pMal-C2, pET, pGEX (Smith *et al.*, Gene 67:31-40, 1988), pMB9 and their derivatives, plasmids such as RP4; phage DNAS, e.g., the numerous derivatives of phage l, e.g., NM989, and other phage DNA, e.g., M13 and filamentous single stranded phage DNA; yeast plasmids such as the 2m plasmid or derivatives thereof; vectors useful in eukaryotic cells, such as vectors useful in insect or mammalian cells; vectors derived from combinations of plasmids and phage DNAs, such as plasmids that have been modified to employ phage DNA or other expression control sequences; and the like.

Yeast expression systems can also be used according to the invention to express hVLR1. For example, the non-fusion pYES2 vector (*Xba*I, *Sph*I, *Sho*I, *Not*I, *Gst*XI, *Eco*RI, *Bst*XI, *Bam*H1, *Sac*I, *Kpn*1, and *Hind*III cloning sit; Invitrogen) or the fusion pYESHisA, B, C (*Xba*I, *Sph*I, *Sho*I, *Not*I, *Bst*XI, *Eco*RI, *Bam*H1, *Sac*I, *Kpn*I, and *Hind*III cloning site, N-terminal peptide purified with ProBond resin and cleaved with enterokinase; Invitrogen), to mention just two, can be employed according to the invention.

Preferred vectors, particularly for cellular assays *in vitro* and *in vivo*, are viral vectors, such as lentiviruses, retroviruses, herpes viruses, adenoviruses,

adeno-associated viruses, vaccinia virus, baculovirus, and other recombinant viruses with desirable cellular tropism. Thus, a gene encoding a functional or mutant protein or polypeptide domain fragment thereof can be introduced *in vivo*, *ex vivo*, or *in vitro* using a viral vector or through direct introduction of DNA. Expression in targeted tissues can be effected by targeting the transgenic vector to specific cells, such as with a viral vector or a receptor ligand, or by using a tissue-specific promoter, or both. Targeted gene delivery is described in International Patent Publication WO 95/28494, published October 1995.

Viral vectors commonly used for *in vivo* or *ex vivo* targeting and therapy procedures are DNA-based vectors and retroviral vectors. Methods for constructing and using viral vectors are known in the art (*see, e.g.*, Miller and Rosman, BioTechniques, 7:980-990, 1992). Preferably, the viral vectors are replication defective, that is, they are unable to replicate autonomously in the target cell. Preferably, the replication defective virus is a minimal virus, *i.e.*, it retains only the sequences of its genome which are necessary for encapsidating the genome to produce viral particles.

DNA viral vectors include an attenuated or defective DNA virus, such as but not limited to herpes simplex virus (HSV), papillomavirus, Epstein Barr virus (EBV), adenovirus, adeno-associated virus (AAV), and the like. Defective viruses, which entirely or almost entirely lack viral genes, are preferred. Defective virus is not infective after introduction into a cell. Use of defective viral vectors allows for administration to cells in a specific, localized area, without concern that the vector can infect other cells. Thus, a specific tissue can be specifically targeted. Examples of particular vectors include, but are not limited to, a defective herpes virus 1 (HSV1) vector (Kaplitt *et al.*, Molec. Cell. Neurosci. 2:320-330, 1991), defective herpes virus vector lacking a glyco-protein L gene (Patent Publication RD 371005 A), or other defective herpes virus vectors (International Patent Publication No. WO 94/21807, published September 29, 1994; International Patent Publication No. WO 92/05263, published April 2, 1994); an attenuated adenovirus vector, such as the vector described by Stratford-Perricaudet *et al.* (J. Clin. Invest. 90:626-630, 1992; see also

La Salle *et al.*, Science 259:988-990, 1993); and a defective adeno-associated virus vector (Samulski *et al.*, J. Virol. 61:3096-3101, 1987; Samulski *et al.*, J. Virol. 63:3822-3828, 1989; Lebkowski *et al.*, Mol. Cell. Biol. 8:3988-3996, 1988).

Various companies produce viral vectors commercially, including but by no means limited to Avigen, Inc. (Alameda, CA; AAV vectors), Cell Genesys (Foster City, CA; retroviral, adenoviral, AAV vectors, and lentiviral vectors), Clontech (retroviral and baculoviral vectors), Genovo, Inc. (Sharon Hill, PA; adenoviral and AAV vectors), Genvec (adenoviral vectors), IntroGene (Leiden, Netherlands; adenoviral vectors), Molecular Medicine (retroviral, adenoviral, AAV, and herpes viral vectors), Norgen (adenoviral vectors), Oxford BioMedica (Oxford, United Kingdom; lentiviral vectors), and Transgene (Strasbourg, France; adenoviral, vaccinia, retroviral, and lentiviral vectors).

Preferably, for *in vivo* administration, an appropriate immunosuppressive treatment is employed in conjunction with the viral vector, *e.g.*, adenovirus vector, to avoid immuno-deactivation of the viral vector and transfected cells. For example, immunosuppressive cytokines, such as interleukin-12 (IL-12), interferon-g (IFN-g), or anti-CD4 antibody, can be administered to block humoral or cellular immune responses to the viral vectors (*see, e.g.*, Wilson, Nature Medicine, 1995). In that regard, it is advantageous to employ a viral vector that is engineered to express a minimal number of antigens.

Adenovirus vectors. Adenoviruses are eukaryotic DNA viruses that can be modified to efficiently deliver a nucleic acid of the invention to a variety of cell types. Various serotypes of adenovirus exist. Of these serotypes, preference is given, within the scope of the present invention, to using type 2 or type 5 human adenoviruses (Ad 2 or Ad 5) or adenoviruses of animal origin (see WO94/26914). Those adenoviruses of animal origin which can be used within the scope of the present invention include adenoviruses of canine, bovine, murine (example: Mav1, Beard *et al.*, Virology 75 (1990) 81), ovine, porcine, avian, and simian (example: SAV) origin. Preferably, the adenovirus of animal origin is a canine adenovirus, more preferably a CAV2 adenovirus (e.g. Manhattan or A26/61 strain (ATCC VR-800), for

example). Various replication defective adenovirus and minimum adenovirus vectors have been described (WO94/26914, WO95/02697, WO94/28938, WO94/28152, WO94/12649, WO95/02697 WO96/22378). The replication defective recombinant adenoviruses according to the invention can be prepared by any technique known to the person skilled in the art (Levrero *et al.*, Gene 101:195 1991; EP 185 573; Graham, EMBO J. 3:2917, 1984; Graham *et al.*, J. Gen. Virol. 36:59 1977). Recombinant adenoviruses are recovered and purified using standard molecular biological techniques, which are well known to one of ordinary skill in the art.

Adeno-associated viruses. The adeno-associated viruses (AAV) are DNA viruses of relatively small size which can integrate, in a stable and site-specific manner, into the genome of the cells which they infect. They are able to infect a wide spectrum of cells without inducing any effects on cellular growth, morphology or differentiation, and they do not appear to be involved in human pathologies. The AAV genome has been cloned, sequenced and characterized. The use of vectors derived from the AAVs for transferring genes *in vitro* and *in vivo* has been described (see WO 91/18088; WO 93/09239; US 4,797,368, US 5,139,941, EP 488 528). The replication defective recombinant AAVs according to the invention can be prepared by cotransfecting a plasmid containing the nucleic acid sequence of interest flanked by two AAV inverted terminal repeat (ITR) regions, and a plasmid carrying the AAV encapsidation genes (rep and cap genes), into a cell line which is infected with a human helper virus (for example an adenovirus). The AAV recombinants which are produced are then purified by standard techniques.

Retrovirus vectors. In another embodiment the gene can be introduced in a retroviral vector, *e.g.*, as described in Anderson *et al.*, U.S. Patent No. 5,399,346; Mann *et al.*, 1983, Cell 33:153; Temin *et al.*, U.S. Patent No. 4,650,764; Temin *et al.*, U.S. Patent No. 4,980,289; Markowitz *et al.*, 1988, J. Virol. 62:1120; Temin *et al.*, U.S. Patent No. 5,124,263; EP 453242, EP178220; Bernstein *et al.* Genet. Eng. 7 (1985) 235; McCormick, BioTechnology 3 (1985) 689; International Patent Publication No. WO 95/07358, published March 16, 1995, by Dougherty *et al.*; and Kuo *et al.*, 1993. Blood 82:845. The retroviruses are integrating viruses which infect

dividing cells. The retrovirus genome includes two LTRs, an encapsidation sequence and three coding regions (gag, pol and env). In recombinant retroviral vectors, the *gag*, *pol* and *env* genes are generally deleted, in whole or in part, and replaced with a heterologous nucleic acid sequence of interest. These vectors can be constructed from different types of retrovirus, such as, HIV, MoMuLV ("murine Moloney leukaemia virus" MSV ("murine Moloney sarcoma virus"), HaSV ("Harvey sarcoma virus"); SNV ("spleen necrosis virus"); RSV ("Rous sarcoma virus") and Friend virus. Suitable packaging cell lines have been described in the prior art, in particular the cell line PA317 (US 4,861,719); the PsiCRIP cell line (WO 90/02806) and the GP+envAm-12 cell line (WO 89/07150). In addition, the recombinant retroviral vectors can contain modifications within the LTRs for suppressing transcriptional activity as well as extensive encapsidation sequences which may include a part of the gag gene (Bender *et al.*, J. Virol. 61:1639, 1987). Recombinant retroviral vectors are purified by standard techniques known to those having ordinary skill in the art.

Retroviral vectors can be constructed to function as infectious particles or to undergo a single round of transfection. In the former case, the virus is modified to retain all of its genes except for those responsible for oncogenic transformation properties, and to express the heterologous gene. Non-infectious viral vectors are manipulated to destroy the viral packaging signal, but retain the structural genes required to package the co-introduced virus engineered to contain the heterologous gene and the packaging signals. Thus, the viral particles that are produced are not capable of producing additional virus.

Retrovirus vectors can also be introduced by DNA viruses, which permits one cycle of retroviral replication and amplifies tranfection efficiency (*see* WO 95/22617, WO 95/26411, WO 96/39036, WO 97/19182).

Lentivirus vectors. In another embodiment, lentiviral vectors are can be used as agents for the direct delivery and sustained expression of a transgene in several tissue types, including brain, retina, muscle, liver and blood. The vectors can efficiently transduce dividing and nondividing cells in these tissues, and maintain long-term expression of the gene of interest. For a review, *see*, Naldini, Curr. Opin.

Biotechnol., 9:457-63, 1998; *see also* Zufferey, *et al.*, J. Virol., 72:9873-80, 1998). Lentiviral packaging cell lines are available and known generally in the art. They facilitate the production of high-titer lentivirus vectors for gene therapy. An example is a tetracycline-inducible VSV-G pseudotyped lentivirus packaging cell line which can generate virusparticles at titers greater than 106 IU/ml for at least 3 to 4 days (Kafri, *et al.*, J. Virol., 73: 576-584, 1999). The vector produced by the inducible cell line can be concentrated as needed for efficiently transducing nondividing cells *in vitro* and *in vivo*.

Non-viral vectors. In another embodiment, the vector can be introduced *in vivo* by lipofection, as naked DNA, or with other transfection facilitating agents (peptides, polymers, etc.). Synthetic cationic lipids can be used to prepare liposomes for in vivo transfection of a gene encoding a marker (Felgner, et. al., Proc. Natl. Acad. Sci. U.S.A. 84:7413-7417, 1987; Felgner and Ringold, Science 337:387-388, 1989; see Mackey, *et al.*, Proc. Natl. Acad. Sci. U.S.A. 85:8027-8031, 1988; Ulmer *et al.*, Science 259:1745-1748, 1993). Useful lipid compounds and compositions for transfer of nucleic acids are described in International Patent Publications WO95/18863 and WO96/17823, and in U.S. Patent No. 5,459,127. Lipids may be chemically coupled to other molecules for the purpose of targeting (see Mackey, et. al., supra). Targeted peptides, *e.g.*, hormones or neurotransmitters, and proteins such as antibodies, or non-peptide molecules could be coupled to liposomes chemically.

Other molecules are also useful for facilitating transfection of a nucleic acid *in vivo*, such as a cationic oligopeptide (*e.g.*, International Patent Publication WO95/21931), peptides derived from DNA binding proteins (*e.g.*, International Patent Publication WO96/25508), or a cationic polymer (*e.g.*, International Patent Publication WO95/21931).

It is also possible to introduce the vector in vivo as a naked DNA plasmid. Naked DNA vectors for gene therapy can be introduced into the desired host cells by methods known in the art, *e.g.*, electroporation, microinjection, cell fusion, DEAE dextran, calcium phosphate precipitation, use of a gene gun, or use of a DNA

vector transporter (see, e.g., Wu *et al.*, J. Biol. Chem. 267:963-967, 1992; Wu and Wu, J. Biol. Chem. 263:14621-14624, 1988; Hartmut *et al.*, Canadian Patent Application No. 2,012,311, filed March 15, 1990; Williams *et al.*, Proc. Natl. Acad. Sci. USA 88:2726-2730, 1991). Receptor-mediated DNA delivery approaches can also be used (Curiel *et al.*, Hum. Gene Ther. 3:147-154, 1992; Wu and Wu, J. Biol. Chem. 262:4429-4432, 1987). US Patent Nos. 5,580,859 and 5,589,466 disclose delivery of exogenous DNA sequences, free of transfection facilitating agents, in a mammal. Recently, a relatively low voltage, high efficiency *in vivo* DNA transfer echnique, termed electrotransfer, has been described (Mir *et al.*, C.P. Acad. Sci., 321:893, 1998; WO 99/01157; WO 99/01158; WO 99/01175).

Antibodies to hVLR1

According to the invention, hVLR1 polypeptides produced recombinantly or by chemical synthesis, and fragments or other derivatives or analogs thereof, including fusion proteins, may be used as an immunogen to generate antibodies that recognize the hVLR1 polypeptide. Such antibodies include but are not limited to polyclonal, monoclonal, chimeric, single chain, Fab fragments, and an Fab expression library. Such an antibody is specific for human hVLR1.

Various procedures known in the art may be used for the production of polyclonal antibodies to hVLR1 polypeptide or derivative or analog thereof. For the production of antibody, various host animals can be immunized by injection with the hVLR1 polypeptide, or a derivative (*e.g.*, fragment or fusion protein) thereof, including but not limited to rabbits, mice, rats, sheep, goats, etc. In one embodiment, the hVLR1 polypeptide or fragment thereof can be conjugated to an immunogenic carrier, *e.g.*, bovine serum albumin (BSA) or keyhole limpet hemocyanin (KLH). Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as BCG (*bacille Calmette-Guerin*) and *Corynebacterium parvum*.

For preparation of monoclonal antibodies directed toward the hVLR1 polypeptide, or fragment, analog, or derivative thereof, any technique that provides for the production of antibody molecules by continuous cell lines in culture may be used. These include but are not limited to the hybridoma technique originally developed by Kohler and Milstein (Nature 256:495-497, 1975), as well as the trioma technique, the human B-cell hybridoma technique (Kozbor *et al.*, Immunology Today 4:72, 1983; Cote *et al.*, Proc. Natl. Acad. Sci. U.S.A. 80:2026-2030, 1983), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole *et al.*, in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96, 1985). Production of human antibodies by CDR grafting is described in U.S. Patent Nos. 5,585,089, 5,693,761, and 5,693,762 to Queen *et al.*, and also in U.S. Patent No. 5,225,539 to Winter and International Patent Application PCT/WO91/09967 by Adau *et al.* In an additional embodiment of the invention, monoclonal antibodies can be produced in germ-free animals (International Patent Publication No. WO 89/12690, published 28 December 1989). In fact, according to the invention, techniques developed for the production of "chimeric antibodies" (Morrison *et al.*, J. Bacteriol. 159:870, 1984); Neuberger *et al.*, Nature 312:604-608, 1984; Takeda *et al.*, Nature 314:452-454, 1985) by splicing the genes from a mouse antibody molecule specific for an hVLR1 polypeptide together with genes from a human antibody molecule of appropriate biological activity can be used; such antibodies are within the scope of this invention. Such human or humanized chimeric antibodies are preferred for use in therapy of human diseases or disorders (described *infra*), since the human or humanized antibodies are much less likely than xenogenic antibodies to induce an immune response, in particular an allergic response, themselves.

According to the invention, techniques described for the production of single chain antibodies (U.S. Patent Nos. 5,476,786 and 5,132,405 to Huston; U.S. Patent 4,946,778) can be adapted to produce hVLR1 polypeptide-specific single chain antibodies. An additional embodiment of the invention utilizes the techniques described for the construction of Fab expression libraries (Huse *et al.*, Science 246:1275-1281, 1989) to allow rapid and easy identification of monoclonal Fab

fragments with the desired specificity for an hVLR1 polypeptide, or its derivatives, or analogs.

Antibody fragments which contain the idiotype of the antibody molecule can be generated by known techniques. For example, such fragments include but are not limited to: the $F(ab')_2$ fragment which can be produced by pepsin digestion of the antibody molecule; the Fab' fragments which can be generated by reducing the disulfide bridges of the $F(ab')_2$ fragment, and the Fab fragments which can be generated by treating the antibody molecule with papain and a reducing agent.

In the production of antibodies, screening for the desired antibody can be accomplished by techniques known in the art, *e.g.*, radioimmunoassay, ELISA (enzyme-linked immunosorbant assay), "sandwich" immunoassays, immunoradiometric assays, gel diffusion precipitin reactions, immunodiffusion assays, *in situ* immunoassays (using colloidal gold, enzyme or radioisotope labels, for example), western blots, precipitation reactions, agglutination assays (*e.g.*, gel agglutination assays, hemagglutination assays), complement fixation assays, immunofluorescence assays, protein A assays, and immunoelectrophoresis assays, etc. In one embodiment, antibody binding is detected by detecting a label on the primary antibody. In another embodiment, the primary antibody is detected by detecting binding of a secondary antibody or reagent to the primary antibody. In a further embodiment, the secondary antibody is labeled. Many means are known in the art for detecting binding in an immunoassay and are within the scope of the present invention. For example, to select antibodies which recognize a specific epitope of an hVLR1 polypeptide, one may assay generated hybridomas for a product which binds to an hVLR1 polypeptide fragment containing such epitope. For selection of an antibody specific to an hVLR1 polypeptide from a particular species of animal, one can select on the basis of positive binding with hVLR1 polypeptide expressed by or isolated from cells of that species of animal.

The foregoing antibodies can be used in methods known in the art relating to the localization and activity of the hVLR1 polypeptide, *e.g.*, for Western blotting, imaging hVLR1 polypeptide *in situ*, measuring levels thereof in appropriate

physiological samples, etc. using any of the detection techniques mentioned above or known in the art. Such antibodies can also be used in assays for ligand binding, *e.g.*, as described in U.S. Patent No. 5,679,582.

In a specific embodiment, antibodies that agonize or antagonize the activity of hVLR1 polypeptide can be generated. Such antibodies can be tested using the assays described *infra* for identifying ligands.

Detection of hVLR1 Expression

One of ordinary skill in the art can use hVLR1-specific oligonucleotides (PCR primers and probes) to detect expression of hVLR1 mRNA. Expression can be detected by Northern analysis or by reverse transcriptase-PCR (RT-PCR). Alternatively, mRNA can be detected by expression of the encoded protein, *e.g.*, in a reticulocyte assay, or by making cDNA and expressing it in a *Xenopous* oocyte assay. However, these latter techniques are more cumbersome and difficult; the Northern or RT-PCR analysis is preferred.

Similarly, one can use antibodies to hVLR1 to detect expression by immunoassay. For example, immunohistology techniques permit detection of expression of hVLR1 receptor. By manipulating the assay conditions, one can distinguish extracellular and intracellular expression. Antibodies for immunodetection of hVLR1 may be specific for hVLR1 or for a tag fused to the hVLR1 protein. In addition, it is possible to use biochemical techniques, such as ligand binding affinity, to establish that the functional hVLR1 protein is present on cells.

For either technique, cells for testing can be obtained from nasal tissues, particularly vomeronasal tissues, from human biopsy or surgical procedures (*e.g.*, rhinoplasties). Alternatively, the cells can be obtained from transgenic animals, and these techniques permit evaluation of hVLR1 expression in the animals. These techniques may also be used to detect hVLR1 expression in tissue culture.

Screening and Chemistry

Identification and isolation of hVLR1 provides for development of screening assays, particularly for high throughput screening of molecules that up- or down-regulate the activity of hVLR1, *e.g.*, by permitting expression of hVLR1 in

quantities greater than can be isolated from natural sources, or in indicator cells that are specially engineered to indicate the activity of hVLR1 expressed after transfection or transformation of the cells. Accordingly, the present invention contemplates methods for identifying specific ligands of hVLR1 using various screening assays known in the art. Furthermore, the invention permits identification of ligands that selectively bind hVLR1 to a greater degree than to other pheromone-like receptors.

Any screening technique known in the art can be used to screen for hVLR1 agonists or antagonists. The present invention contemplates screens for small molecule ligands or ligand analogs and mimics, as well as screens for natural ligands that bind to and agonize or antagonize the activity of hVLR1 *in vivo*. For example, natural products libraries can be screened using assays of the invention for molecules that agonize or antagonize hVLR1 activity. Generally, compounds are tested for the ability to compete with labeled pheromone-like substrate or a pheromone-like analog, for binding to the hVLR1. Screens can either be "cell-free", *i.e.*, binding assays of receptor protein with candidate compounds, where the protein is on a solid support in a liposome or micelle, or cell-based, in which the protein is found in a cell membrane.

It is also possible to directly label the test compound to evaluate binding.

Knowledge of the primary sequence of the protein, and the similarity of that sequence with proteins of known function, can provide an initial clue as to the inhibitors or antagonists of the protein. Identification and screening of antagonists is further facilitated by determining structural features of the protein, *e.g.*, using X-ray crystallography, neutron diffraction, nuclear magnetic resonance spectrometry, and other techniques for structure determination. These techniques provide for the rational design or identification of agonists and antagonists.

Another approach uses recombinant bacteriophage to produce large libraries. Using the "phage method" (Scott and Smith, Science 249:386-390, 1990; Cwirla, *et al.*, Proc. Natl. Acad. Sci., 87:6378-6382, 1990; Devlin *et al.*, Science, 49:404-406, 1990), very large libraries can be constructed (10^6-10^8 chemical entities). A second approach uses primarily chemical methods, of which the Geysen method

(Geysen *et al.*, Molecular Immunology 23:709-715, 1986; Geysen *et al.* J. Immunologic Method 102:259-274, 1987; and the method of Fodor *et al.* (Science 251:767-773, 1991) are examples. Furka *et al.* (14th International Congress of Biochemistry, Volume #5, Abstract FR:013, 1988; Furka, Int. J. Peptide Protein Res. 37:487-493, 1991), Houghton (U.S. Patent No. 4,631,211, issued December 1986) and Rutter *et al.* (U.S. Patent No. 5,010,175, issued April 23, 1991) describe methods to produce a mixture of peptides that can be tested as agonists or antagonists.

In another aspect, synthetic libraries (Needels *et al.*, Proc. Natl. Acad. Sci. USA 90:10700-4, 1993; Ohlmeyer *et al.*, Proc. Natl. Acad. Sci. USA 90:10922-10926, 1993; Lam *et al.*, International Patent Publication No. WO 92/00252; Kocis *et al.*, International Patent Publication No. WO 9428028) and the like can be used to screen for hVLR1 ligands according to the present invention.

In another embodiment, a yeast screening assay, useful for testing agonists and antagonists of mammalian G-protein coupled receptors, *e.g.*, as disclosed in U.S. Patent No. 5,482,832, can be used.

Test compounds are screened from large libraries of synthetic or natural compounds. Numerous means are currently used for random and directed synthesis of saccharide, peptide, and nucleic acid based compounds. Synthetic compound libraries are commercially available from Maybridge Chemical Co. (Trevillet, Cornwall, UK), Comgenex (Princeton, NJ), Brandon Associates (Merrimack, NH), and Microsource (New Milford, CT). A rare chemical library is available from Aldrich (Milwaukee, WI). Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available from *e.g.* Pan Laboratories (Bothell, WA) or MycoSearch (NC), or are readily producible. Additionally, natural and synthetically produced libraries and compounds are readily modified through conventional chemical, physical, and biochemical means (Blondelle *et al.*, Tib Tech, 14:60, 1996).

In vivo *and* ex vivo *screening methods*

Intact cells or whole animals expressing a gene encoding hVLR1 can be used in screening methods to identify candidate drugs.

In one series of embodiments, a permanent cell line is established. Alternatively, cells (including without limitation mammalian, insect, yeast, or bacterial cells) are transiently programmed to express an hVLR1 gene by introduction of appropriate DNA or mRNA. Identification of candidate compounds can be achieved using any suitable assay, including without limitation (i) assays that measure selective binding of test compounds to hVLR1 (ii) assays that measure the ability of a test compound to modify (*i.e.*, inhibit or enhance) a measurable activity or function of hVLR1 and (iii) assays that measure the ability of a compound to modify (*i.e.*, inhibit or enhance) the transcriptional activity of sequences derived from the promoter (*i.e.*, regulatory) regions the hVLR1 gene.

Identification of ligands for olfactory receptors can be achieved by screening for functional expression of receptors, *e.g. see* Krautwurst *et al.*, Cell 95:917-926, 1998.

Gene targeting technology to introduce mutations in putative pheromone receptor genes. By generating alleles differentially tagged with the histological markers, the putative pheromone receptor gene expression pattern can be detected topographically. These histological markers combines the intrinsic fluorescent properties of green fluorescent protein (GFP) with the microtubule-binding properties of the tau protein, assuring that the fluorescent label will be exported down the axons to the axon terminals of the vomeronasal sensory neurons that express the marker-tagged receptor genes. A deletion allele of the putative receptor gene determines whether expression of the putative receptor gene is a determinant of the pattern of axonal projections. The intronless coding region is excised from start to stop codon and replaced with a GFP-IRES-taulacZ cassette. To determine whether other seven-transmembrane proteins can substitute for the putative receptor gene, the coding region of the putative receptor gene can be replaced with that of a known odorant receptor gene, thus generating a swap. The consequence of

this swap on the topography of the projections can be examined by visualizing the taulacZ expression (Rodriquez *et al.*, Cell 97:199-208, 1999).

Transgenic mammals can be prepared for evaluating the molecular mechanisms of hVLR1. Such mammals provide excellent models for screening or testing drug candidates. Thus, hVLR1 "knock-in" mammals can be prepared for evaluating the molecular biology of this system in greater detail than is possible with human subjects. It is also possible to evaluate compounds or diseases on "knockout" animals, *e.g.*, to identify a compound that can compensate for a defect in hVLR1. Both technologies permit manipulation of single units of genetic information in their natural position in a cell genome and to examine the results of that manipulation in the background of a terminally differentiated organism.

A "knockin" mammal is a mammal in which an endogenous gene is substituted with a heterologous gene (Roemer *et al.*, New Biol. 3:331, 1991). Preferably, the heterologous gene is "knocked-in" to a locus of interest, either the subject of evaluation(in which case the gene may be a reporter gene; *see* Elefanty *et al.*, Proc Natl Acad Sci USA 95:11897,1998) of expression or function of a homologous gene, thereby linking the heterologous gene expression to transcription from the appropriate promoter. This can be achieved by homologous recombination, transposon (Westphal and Leder, Curr Biol 7:530, 1997), using mutant recombination sites (Araki *et al.*, Nucleic Acids Res 25:868, 1997) or PCR (Zhang and Henderson, Biotechniques 25:784, 1998).

A "knockout mammal" is an mammal (*e.g.*, mouse) that contains within its genome a specific gene that has been inactivated by the method of gene targeting (*see, e.g.*, US Patents No. 5,777,195 and No. 5,616,491). A knockout mammal includes both a heterozygote knockout (*i.e.*, one defective allele and one wild-type allele) and a homozygous mutant (*i.e.*, two defective alleles). Preparation of a knockout mammal requires first introducing a nucleic acid construct that will be used to suppress expression of a particular gene into an undifferentiated cell type termed an embryonic stem cell. This cell is then injected into a mammalian embryo. A mammalian embryo with an integrated cell is then implanted into a foster mother

for the duration of gestation. Zhou, *et al.* (Genes and Development, 9:2623-34, 1995) describes PPCA knock-out mice.

The term "knockout" refers to partial or complete suppression of the expression of at least a portion of a protein encoded by an endogenous DNA sequence in a cell. The term "knockout construct" refers to a nucleic acid sequence that is designed to decrease or suppress expression of a protein encoded by endogenous DNA sequences in a cell. The nucleic acid sequence used as the knockout construct is typically comprised of (1) DNA from some portion of the gene (exon sequence, intron sequence, and/or promoter sequence) to be suppressed and (2) a marker sequence used to detect the presence of the knockout construct in the cell. The knockout construct is inserted into a cell, and integrates with the genomic DNA of the cell in such a position so as to prevent or interrupt transcription of the native DNA sequence. Such insertion usually occurs by homologous recombination (*i.e.*, regions of the knockout construct that are homologous to endogenous DNA sequences hybridize to each other when the knockout construct is inserted into the cell and recombine so that the knockout construct is incorporated into the corresponding position of the endogenous DNA). The knockout construct nucleic acid sequence may comprise 1) a full or partial sequence of one or more exons and/or introns of the gene to be suppressed, 2) a full or partial promoter sequence of the gene to be suppressed, or 3) combinations thereof. Typically, the knockout construct is inserted into an embryonic stem cell (ES cell) and is integrated into the ES cell genomic DNA, usually by the process of homologous recombination. This ES cell is then injected into, and integrates with, the developing embryo.

The phrases "disruption of the gene" and "gene disruption" refer to insertion of a nucleic acid sequence into one region of the native DNA sequence (usually one or more exons) and/or the promoter region of a gene so as to decrease or prevent expression of that gene in the cell as compared to the wild-type or naturally occurring sequence of the gene. By way of example, a nucleic acid construct can be prepared containing a DNA sequence encoding an antibiotic resistance gene which is inserted into the DNA sequence that is complementary to the DNA sequence

(promoter and/or coding region) to be disrupted. When this nucleic acid construct is then transfected into a cell, the construct will integrate into the genomic DNA. Thus, many progeny of the cell will no longer express the gene at least in some cells, or will express it at a decreased level, as the DNA is now disrupted by the antibiotic resistance gene.

Generally, the DNA will be at least about 1 kilobase (kb) in length and preferably 3-4 kb in length, thereby providing sufficient complementary sequence for recombination when the knockout construct is introduced into the genomic DNA of the ES cell (discussed below).

Included within the scope of this invention is a mammal in which two or more genes have been knocked out. Such mammals can be generated by repeating the procedures set forth herein for generating each knockout construct, or by breeding to mammals, each with a single gene knocked out, to each other, and screening for those with the double knockout genotype.

Regulated knockout animals can be prepared using various systems, such as the tet-repressor system (*see* US Patent No. 5,654,168) or the Cre-Lox system (*see* US Patents No. 4,959,317 and No. 5,801,030).

In another series of embodiments, transgenic animals are created in which (i) a human hVLR1 is stably inserted into the genome of the transgenic animal; and/or (ii) the endogenous vomeronasal receptor genes are inactivated and replaced with human hVLR1 genes. See, *e.g.*, Coffman, Semin. Nephrol. 17:404, 1997; Esther *et al.*, Lab. Invest. 74:953, 1996; Murakami *et al.*, Blood Press. Suppl. 2:36, 1996. Such animals can be treated with candidate compounds and monitored for responses associated with pheromone stimulation.

High-Throughput Screen

Agents according to the invention may be identified by screening in high-throughput assays, including without limitation cell-based or cell-free assays. It will be appreciated by those skilled in the art that different types of assays can be used to detect different types of agents. Several methods of automated assays have been developed in recent years so as to permit screening of tens of thousands of compounds

in a short period of time, such as assays based on protein stability when contacted with a candidate ligand (see U.S. Patent Nos. 5,679,582 and 5,585,277). Such high-throughput screening methods are particularly preferred. The use of high-throughput screening assays to test for agents is greatly facilitated by the availability of large amounts of purified polypeptides, as provided by the invention.

Preferred Screening Methods

There are several screening methods available for the discovery of non-peptide antagonists. These screens include radio ligand binding, signal transduction, yeast expression, reporter assays, and structure function of existing peptide agonists and non-peptide antagonists. The hVLR1 expressed in yeast displays pharmacological properties similar to that observed for this receptor when expressed in mammalian cells. The utilization of yeast as a screening tool can accelerate the search for novel pheromone analogs. This technology can be utilized for screening of novel compounds that are identified in high throughput screens.

Signal Transduction Assays

G protein coupled receptors (GPCR) are coupled to a variety of heterotrimeric G proteins, which are comprised of α, β, and γ subunits. Upon agonist binding to a GPCR at the cell surface, conformational changes occur within the agonist:GPCR complex, which lead to the dissociation of the G protein α subunit from the βγ subunits. The G_{α} and $G_{\beta\gamma}$ subunits then stimulate a variety of intracellular effectors, which transduce the extracellular signal to the inside of the cell. Various signal transduction systems known to be coupled to GPCRs include adenylate cyclase, phospholipase C, phospholipase A_2, sodium/hydrogen exchange, etc. Thus, measurements of intracellular calcium concentrations and adenylate cyclase activity indicate whether a hit or test compound is functionally behaving as an agonist or antagonist at the vomeronasal-like receptor.

In a specific embodiment, G-protein signal transduction is coupled to expression of a reporter gene, thus permitting a reporter gene screening assay.

EXAMPLES

The present invention will be better understood by reference to the following Examples, which are provided as exemplary of the invention, and not by way of limitation.

EXAMPLE 1: Isolation of Human Vomeronasal-like Receptor Sequences

In an attempt to clone vomeronasal-like receptors from human, we designed degenerate primers specific for selected conserved regions in mouse and rat vomeronasal receptor (VR) genes. This approach was chosen to favor the amplification of VR-like sequences with intact open reading frames (ORFs), reasoning that pseudogenes would differ at these conserved motifs more than functional genes. Five conserved regions were chosen and 8 primers were synthesized, allowing for 12 different primer pairs to be tested.

The sense primers are:

TAC32: CTI AGY CCI AGR AGY TCI TG	(SEQ ID NO:5)
TAC33: ATM GCI ACI CCI AAY TR AC	(SEQ ID NO:6)
TAC145: AAR GCI TCI CCI GAR CAR AGR GCI AC	(SEQ ID NO:7)

The reverse primers are:

TAC35: ARI ARI GCI ACC ATR TAI C	(SEQ ID NO:8)
TAC36: CKI GTI GCY CTY TGY TCI GG	(SEQ ID NO:9)
TAC143: ACR AAI GGR CTI ACI GTI GCR TA	(SEQ ID NO:10)
TAC143': TCR GGI AAR CAI TAD WSI TG	(SEQ ID NO:11)
TAC144: ARI ATI GTY CTI GTI GCY CTY TG.	(SEQ ID NO:12)

Although these primers amplify multiple mouse VR genes, a single band is observed when the PCR product is separated on an agarose gel, because all VR coding sequences are contained in a single exon and have a similar size.

Human genomic DNA from a Caucasian male was then used as template. PCR conditions were: 94°C for 1 min, 48°C for 3 min, 72°C for 3 min, 39 cycles with 6 sec extension per cycle. Amplified PCR products, which were similar in size to those obtained from mouse genomic DNA, were isolated and cloned into

Calcium Mobilization Assay

Whole cells expressing the vomeronasal-like receptor are loaded with a fluorescent dye that chelates calcium ions, such as FURA-2. Upon addition of pheromone-like substrate to these cells, pheromone-like substrate binds to the vomeronasal-like receptors and calcium is released from the intracellular stores. The dye chelates these calcium ions. Spectrophotometric determination of the ratio for dye:calcium complexes to free dye determine the changes in intracellular calcium concentrations upon addition of pheromone-like substrate. Hits from screens and other test compounds can be similarly tested in this assay to functionally characterize them as agonists or antagonists. Increases in intracellular calcium concentrations are expected for compounds with agonist activity while compounds with antagonist activity are expected to block pheromone-like substrate stimulated increases in intracellular calcium concentrations.

Cyclic AMP Accumulation Assay

Upon agonist binding, G_s coupled GPCRs stimulate adenylate cyclase. Adenylate cyclase catalyzes the production of cyclic AMP from adenosine-5'-triphosphate which, in turn, activates protein kinases. G_i coupled GPCRs are also coupled to adenylate cyclase, however, agonist binding to these receptors results in the inhibition of adenylate cyclase and the subsequent inhibition of cAMP. To measure the inhibition of cAMP accumulation, cells expressing G_i coupled receptors must first be stimulated to elevate cAMP levels. This is achieved by treating the cells with forskolin, a diterpene that directly stimulates cAMP production. Co-incubation of cells expressing G_i coupled receptors with forskolin and a functional agonist will result in the inhibition of forskolin-stimulated cAMP accumulation. For a cAMP assay, whole cells stably expressing hVLR1 can be incubated with a test compound, and with forskolin plus a test compound. The cells are then lysed and cAMP levels are measured using the [^{125}I]cAMP scintillation proximity assay (SPA). Functional agonists of G_s coupled receptors are expected to increase cAMP levels above basal levels whereas functional agonists of G_i coupled receptors are expected to inhibit the forskolin-stimulated cAMP accumulation.

Subsequent hybridizations were performed with a probe corresponding to the hVLR1 mRNA sequence flanked by TAC588 and TAC597, which includes part of the CDS. For the evaluation of the expression level of hVLR1 mRNA, a human multiple tissue expression array (MTe, Clonetech) was employed.

hVLR1 mRNA expression was found consistently found in the olfactory mucosa. In contrast, mRNA expression was detected with poor reproducibility in the brain, lung and kidney, probably reflecting very low expression levels in these tissues. (Due to its very high sensitivity, this method does not provide more than qualitative information regarding the levels of hVLR1 mRNA in a given tissue.)

Potential expression of hVLR1 mRNA in non-olfactory tissues was further investigated by dot blot analysis, with a commercially available panel of 76 human polyA$^+$ mRNAs pools, each corresponding to a different tissue. The hVLR1 probe did not detectably hybridize to any sample tested except for a very weak signal with adult cerebellum mRNA (data not shown). The variable presence of hVLR1 transcripts outside the olfactory system is not surprising as in vertebrates, expression of odorant receptor mRNAs has also been reported in non-olfactory tissues such as testis (Parmentier, M. et al. Nature 355, 453-455,1992), heart Drutel, G. et al. Receptors Channels 3, 33-40, 1995), spleen (Blache, P., et al., D. Biochem. Biophys. Res. Commun. 242, 669-672, 1998) and notochord (Nef, S.,et al., Mech. Dev. 55, 65-77,1996); the significance of this ectopic expression is unclear.

Redundant samples of olfactory mucosa from three adult patients were removed during elective surgery for septodermoplasty, following approval of the Institutional Review Board at The Rockefeller University and of the Yale University School of Medicine. The indication for surgery was recurrent nasal bleeds. Tissue was frozen at -70°C within two minutes of having been surgically resected. Human olfactory epithelium is not a uniform sensory sheet but exhibits an irregular and patchy distribution of olfactory sensory neurons within the nasal cavity.(Morrison, E.E. & Costanzo, R.M. J. Comp. Neurol. 297, 1-13, 1990; Morrison, E.E. & Costanzo, R.M. Microsc. Res. Tech. 23, 49-61, 1992). The presence of olfactory

swapped receptor are analyzed. The neurons expressing the human receptor are also easily detected as they will express green flourescent protein (GFP). This will allow the study of hVLR1 ligands by following single cell events, *e.g.*, calcium fluxes.

EXAMPLE 4: hVLR1 Polymorphic Variants

In order to investigate the presence of hVLR1 orthologs in non-human primates, a V1LR1 ortholog was cloned from chimpanzee genomic DNA.

Human genomic DNAs from different ethnic backgrounds were obtained from the Coriell Cell Repositories (New Jersey). Genomic DNA from chimpanzee was extracted from blood obtained from the Yerkes Regional Primate Center (Atlanta, Georgia). Primers used for the amplification of the human and the chimpanzee sequences flanked the V1LR1 ORF, and are as follows: TAC558: TTC TCT GCA GTT GGA CAC ACA AGC (SEQ ID NO:13), and TAC559: GCA AGA GTT ATG ATA AAT AGC TG (SEQ ID NO:14).

The deduced amino acid sequences from two human V1RL1 splice variants (a and b) were aligned with the putative chimpanzee ortholog (cV1RL1 or cVRL1) (SEQ ID NO:17) and the two mouse V1R sequences, mVR23 (SEQ ID NO:18) and mpr2 (SEQ ID NO:19) (Genbank accession number Y12724). The cloned V1LR1 ortholog coding sequence (CDS) codes for a protein sharing 93% identity with VLR1, diverging mainly at the C-terminus (Fig. 7). Again, the 11 V1R-conserved residues and potential N-linked glycosylation site are present within cVLR1.

EXAMPLE 5: Tissue Specificity Of hVLR1 Expression

In order to determine the tissue-specificity of hVLR1 expression, a wide range of tissues and organs were screened by RT-PCR analysis by hybridization with a V1LR1 probe (Fig. 8).

Primers used for the amplification of hVLR1 cDNA in different human tissue samples (Human Multiple Tissue cDNA Panels, Clontach) were located in hVLR1 exons 2 and 4: TAC588: GTT CCC ATG AAC TCA GAA G (SEQ ID NO:13) and TAC597: TGG CTG AGA ATC AAG TCC GT (SEQ ID NO:14).

EXAMPLE 2: Detection of hVLR1 In Human Vomeronasal Tissue

In order to explore the expression of the BH33 gene in the human olfactory system, sensory olfactory epithelium was extracted from patients, RNA isolated from the tissues, and cDNA generated from this RNA. As expected, human olfactory epithelium expressed hVLR1 mRNAs.

Using specific primers on the BH33 coding sequence and an "adapter oligonucleotide", the 5' untranslated sequence of BH33 was cloned by 5'RACE. The sequence of the 5'untranslated region showed, when compared to the genomic sequence of the BH33 locus, that the BH33 gene consists of three exons (see Figure 5B) and that the coding sequence is entirely included in the last exon. The predicted open reading frame would encode for a 313 amino acids seven-transmembrane protein (see Figure 5B).

These data indicated that, in addition of having a full length vomeronasal receptor-like open reading frame, *BH33* has functional splice donor and acceptor sites and that the promoter is also intact. It also demonstrates that the BH33 gene is expressed in human olfactory epithelium.

Another BH33 splice variant was also found to be expressed in the olfactory epithelium, which seemed to use a cryptic splice site inside the BH33 coding sequence, and which used a fourth exon (see Figure 5B). The potential protein product of this splice variant would have 243 amino acids and would share homology to vomeronasal receptors only on the two last thirds of its sequence.

EXAMPLE 3: Knock-In of hVLR1

In order to explore the potential of the hVLR1 to substitute for a mouse vomeronasal receptor, the coding sequence of the mouse VR2 receptor is substituted by the coding sequence of hVLR1 and green fluorescent protein under control of an internal ribosome expression site (IRES), as was done previously with another receptor system (Rodriguez *et al.*, Cell, 1999, 97:199-208). This procedure is shown schematically in Figure 6. The projection pattern of the neurons expressing the

pGEMT. Three pairs of primers allowed the amplification of potential human VR receptor genes, and 5 clones for each of the amplified products were subcloned and sequenced. The sequenced products corresponding to potential human VRs were then used as templates to synthesize radioactive probes for screening a human genomic BAC library.

In a parallel approach, a pair of degenerate primers was used to amplify mouse genomic DNA. 23 different mouse VR receptor partial sequences were cloned and sequenced from this amplified product. This heterogenous product was then used as a template for a probe to screen the human BAC library, at medium stringency (59°C).

From both approaches, 27 human BAC clones were selected, from which 8 distinct subclones containing sequences homologous to mouse or rat VR genes were isolated. One human sequence, termed Bh33, has a complete ORF potentially encoding for a protein similar to the mouse or rat VRs from the VR1 family (Figure 1). Subsequently, it was determined that hVLR1 might have an alternative expression form using an upstream ATG start site in reading frame with the putative start site shown in Figure 1. This "long" form is shown in Figure 2. The Bh33 deduced amino acid sequence is 28% identical and 47% similar to the mouse VR2 sequence, and many of the conserved amino acid residues in rat and mouse VRs are also conserved in Bh33 (see Figure 3).

The seven other BAC subclones were human pseudogenes, as they contained multiple frameshifts and stop codons in the coding sequence (Figure 4).

To find potential polymorphisms in humans, the Bh33 coding sequence was cloned and sequenced from four Caucasian males and individuals of Indonesian, Pygmy, Amerindian, Cambodian, Japanese, Ami, and Adygei extractions. The sequences from the four Caucasians were identical, and two single nucleotide substitutions were found in the Indonesian, Pygmy, Japanese, Cambodian, and Ami subjects, resulting in two amino acid differences (S201 → F, A229 → D). The Amerindian and Adygei subjects each had the S201 → F substitution. None of the allelic variants were found to be interrupted by a stop codon or to contain a frameshift.

WE CLAIM:

1. An isolated human vomeronasal-like receptor.

2. The receptor according to claim 1 which comprises an amino acid sequence having greater than about 90% amino acid sequence identity to the amino acid sequence depicted in SEQ ID NO:2 or SEQ ID NO:4.

3. The receptor according to claim 2 comprising a modification of the amino acid sequence selected from the group consisting of Ser201 → Phe, Ala229 → Asp, and both.

4. The receptor according to claim 1 which is encoded by a DNA molecule having a nucleotide sequence as depicted in SEQ ID NO:1 or SEQ ID NO:3, or SEQ ID NO:17.

5. An antigenic fragment of the receptor of claim 1.

6. A chimeric polypeptide comprising an amino acid sequence of a human vomeronasal-like receptor fused to a heterologous amino acid sequence.

7. The chimeric polypeptide of claim 6, wherein the amino acid sequence having functional activity is selected from a group consisting of a signal peptide, an antibody tag, an expression tag, a chromatographic tag, a cytoplasmic signal domain, and a G-protein binding domain.

8. A nucleic acid encoding the polypeptide of claim 6.

9. A nucleic acid encoding the polypeptide of claim 7.

epithelium within some biopsies was verified with immunohistochemistry for neuron-specific enolase. RNA as extracted from the snap-frozen biopsies using standard protocols. For 5'RACE, reverse transcription and further cycles of amplification were carried out with the SMART kit (Clontech) using reverse nested primers, on RNA isolated from olfactory mucosa.

Our results indicate, contrary to preliminary data (Dulac, C. & Axel, R. Cell 83, 195-206, 1995), that the human genome contains at least one V1R-like gene that is not a pseudogene; furthermore, that this gene is transcribed into a spliced mRNA within cells of the olfactory mucosa of adults; and finally, that the transcript has the potential to encode a seven-transmembrane protein homologous to putative pheromone receptors of rodents.

* * * * * * * * *

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and the accompanying figures. Such modifications are intended to fall within the scope of the appended claims.

It is further to be understood that all values, are approximate, and are provided for description.

All patents, patent applications, publications, and other materials cited herein are hereby incorporated herein reference in their entireties.

20. The vector of claim 19 wherein the nucleic acid encoding a human vomeronasal-like receptor is operatively associated with an expression control sequence that permits expression of the receptor in a host cell.

21. A host cell comprising the vector of claim 20.

22. A method for producing a human vomeronasal-like receptor, which method comprises culturing the host cell of claim 21 under conditions that permit expression of the receptor.

23. An isolated host cell that expresses a human vomeronasal-like receptor, with the proviso that the cell is not a human cell that endogenously express the receptor.

24. A non-human animal that expresses a human vomeronasal-like receptor.

25. An antibody that specifically binds to a human vomeronasal-like receptor.

26. A method of identifying a compound that binds to the receptor of claim 1, which method comprises detecting association of a candidate compound with the receptor when the compound and the receptor, wherein detection of such association indicates that the compound binds to the receptor.

27. The method according to claim 26, wherein the compound is labeled.

28. The method according to claim 26, wherein the receptor is present in a membrane of a cell.

10. A nucleic acid encoding a human vomeronasal-like receptor.

11. The nucleic acid of claim 10 which hybridizes under high stringency conditions to a nucleic acid having a sequence corresponding to or complementary to a nucleic acid sequence as depicted in SEQ ID NO:1 or SEQ ID NO:3.

12. The nucleic acid of claim 11 which encodes a protein comprising an amino acid sequence as depicted in SEQ ID NO:2 or SEQ ID NO:4.

13. The nucleic acid of claim 10 which comprises a nucleotide sequence as depicted in SEQ ID NO:1 or SEQ ID NO:3.

14. An isolated nucleic acid comprising a nucleotide sequence corresponding to or complementary to at least a ten base length of a nucleotide sequence as depicted in SEQ ID NO:1 or SEQ ID NO:3.

15. The nucleic acid of claim 14 which is single stranded.

16. The nucleic acid of claim 15 which is selected from the group consisting of SEQ ID NOS:5-14.

17. The nucleic acid of claim 15 which is labeled.

18. The nucleic acid of claim 15 which hybridizes under intracellular conditions to an mRNA encoding a human vomeronasal-like receptor.

19. A vector comprising the nucleic acid encoding a human vomeronasal-like receptor of claim 10.

37. The receptor according to claim 36 which comprises an amino acid sequence having greater than about 90% amino acid sequence identity to the amino acid sequence depicted in SEQ ID NO:2 or SEQ ID NO:4.

38. The receptor according to claim 37 comprising a modification of the amino acid sequence selected from the group consisting of Ser201 → Phe, Ala229 → Asp, and both.

39. The receptor according to claim 36 which is encoded by a DNA molecule having a nucleotide sequence as depicted in SEQ ID NO:1 or SEQ ID NO:3.

40. A nucleic acid encoding a primate vomeronasal-like receptor.

41. The nucleic acid of claim 40 which hybridizes under high stringency conditions to a nucleic acid having a sequence corresponding to or complementary to a nucleic acid sequence as depicted in SEQ ID NO:1 or SEQ ID NO:3.

42. The nucleic acid of claim 41 which encodes a protein comprising an amino acid sequence as depicted in SEQ ID NO:2 or SEQ ID NO:4.

43. The nucleic acid of claim 40 which comprises a nucleotide sequence as depicted in SEQ ID NO:1 or SEQ ID NO:3.

29. The method according to claim 26, wherein the receptor is present free of cellular components.

30. The method according to claim 26, wherein the compound that binds to the receptor modulates receptor signaling.

31. A method for detecting a compound that agonizes the receptor of claim 1, which method comprises detecting G-protein activation in a cell that expresses the receptor when the cell is contacted with a compound that binds to the receptor.

32. A method for detecting expression of a human vomeronasal-like receptor in a cell, which method comprises detecting the presence of mRNA encoding the receptor in the cell.

33. A method for detecting expression of a human vomeronasal-like receptor in a cell, which method comprises detecting the presence of a human vomeronasal-like receptor in the cell.

34. A method for identifying an allelic variant of a gene encoding a human vomeronasal-like receptor, which method comprises detecting a polymorphism in a gene encoding a human vomeronasal-like receptor when compared to a sequence of a gene encoding a human vomeronasal-like receptor as depicted in SEQ ID NO:1.

35. The method according to claim 34, wherein the polymorphism is detected by sequencing.

36. An isolated primate vomeronasal-like receptor.

FIG. 2a

1/1 31/11
atg gtt gga gac aca tta aaa ctt ctg tct cca ctg atg aca aga tac ttc ttt ctg ctt
M V G D T L K L L S P L M T R Y F F L L
61/21 91/31
ttt tat tct act gat tct tca gac ctc aat gaa aat caa cat ccc cta gat ttt gat gaa
F Y S T D S S D L N E N Q H P L D F D E
121/41 151/51
atg gct ttt gga aaa gta aaa tca ggg att agc ttc ctc att cag act gga gtt ggg atc
M A F G K V K S G I S F L I Q T G V G I
181/61 211/71
ctg gga aat tcc ttt ctc ctt tgt ttt tat aac tta att ttg ttc act gga cac aag ctg
L G N S F L L C F Y N L I L F T G H K L
241/81 271/91
aga ccc acg gac ttg att ctc agc caa ctg gcc ttg gct aac tcc atg gtc ctt ttc ttt
R P T D L I L S Q L A L A N S M V L F F
301/101 331/111
aaa ggg ata cct cag aca atg gca gct ttt gga ttg aaa tat ttg ctg aat gac act gga
K G I P Q T M A A F G L K Y L L N D T G
361/121 391/131
tgt aag ttt gtc ttt tat tat cac agg gtg ggc aca aga gtt tcc ctc agc acc atc tgc
C K F V F Y Y H R V G T R V S L S T I C
421/141 451/151
ctt ctc aat gga ttc caa gcc att aag ctc aac ccc agt ata tgc agg tgg atg gag atc
L L N G F Q A I K L N P S I C R W M E I
481/161 511/171
aag att aga tcc cca agg ttt att gac ttc tgt tgt ctc ctc tgc tgg gcc ccc cat gtc
K I R S P R F I D F C C L L C W A P H V
541/181 571/191
ttg atg aat gca tct gtt ctt cta tta gtg aat ggc cca ctg aat agc aaa aac agt agt
L M N A S V L L L V N G P L N S K N S S
601/201 631/211
gca aaa aac aat tat gga tac tgt tct tac aaa gca tca aag aga ttt agc tca tta cat
A K N N Y G Y C S Y K A S K R F S S L H

A——————A

FIG. 1

1/1 31/11
atg gct ttt gga aaa gta aaa tca ggg att agc ttc ctc att cag act gga gtt ggg atc
M A F G K V K S G I S F L I Q T G V G I
61/21 91/31
ctg gga aat tcc ttt ctc ctt tgt ttt tat aac tta att ttg ttc act gga cac aag ctg
L G N S F L L C F Y N L I L F T G H K L
121/41 151/51
aga ccc acg gac ttg att ctc agc caa ctg gcc ttg gct aac tcc atg gtc ctt ttc ttt
R P T D L I L S Q L A L A N S M V L F F
181/61 211/71
aaa ggg ata cct cag aca atg gca gct ttt gga ttg aaa tat ttg ctg aat gac act gga
K G I P Q T M A A F G L K Y L L N D T G
241/81 271/91
tgt aag ttt gtc ttt tat tat cac agg gtg ggc aca aga gtt tcc ctc agc acc atc tgc
C K F V F Y Y H R V G T R V S L S T I C
301/101 331/111
ctt ctc aat gga ttc caa gcc att aag ctc aac ccc agt ata tgc agg tgg atg gag atc
L L N G F Q A I K L N P S I C R W M E I
361/121 391/131
aag att aga tcc cca agg ttt att gac ttc tgt tgt ctc ctc tgc tgg gcc ccc cat gtc
K I R S P R F I D F C C L L C W A P H V
421/141 451/151
ttg atg aat gca tct gtt ctt cta tta gtg aat ggc cca ctg aat agc aaa aac agt agt
L M N A S V L L L V N G P L N S K N S S
481/161 511/171
gca aaa aac aat tat gga tac tgt tct tac aaa gca tca aag aga ttt agc tca tta cat
A K N N Y G Y C S Y K A S K R F S S L H
541/181 571/191
gca gtc tta tat ttt tcc cct gat ttt atg agt ttg ggc ttc atg gtc tgg gcc agt ggc
A V L Y F S P D F M S L G F M V W A S G
601/201 631/211
tcc atg gtc ttc ttc ctc tac aga cac aag cag caa gtc caa cac aat cac agc aac aga
S M V F F L Y R H K Q Q V Q H N H S N R
661/221 691/231
ctc tcc tgc aga cct tcc cag gaa gcc aga gcc aca cac acc atc atg gtc ctg gtg agc
L S C R P S Q E A R A T H T I M V L V S
721/241 751/251
tcc ttt ttt gtt ttc tat tca gtc cat agt ttt ctg aca att tgg aca act gta gtt gca
S F F V F Y S V H S F L T I W T T V V A
781/261 811/271
aac cca ggc cag tgg ata gtg acc aac tct gtg ttg gtc gcc tca tgt ttc cca gca cgc
N P G Q W I V T N S V L V A S C F P A R
841/281 871/291
agc cct ttt gtc ctc atc atg agt gat act cat atc tct cag ttc tgt ttt gcc tgc agg
S P F V L I M S D T H I S Q F C F A C R
901/301 931/311
aca agg aaa aca ctc ttt cct aat ctg gtt gtc atg cca tga
T R K T L F P N L V V M P *

FIG. 3a

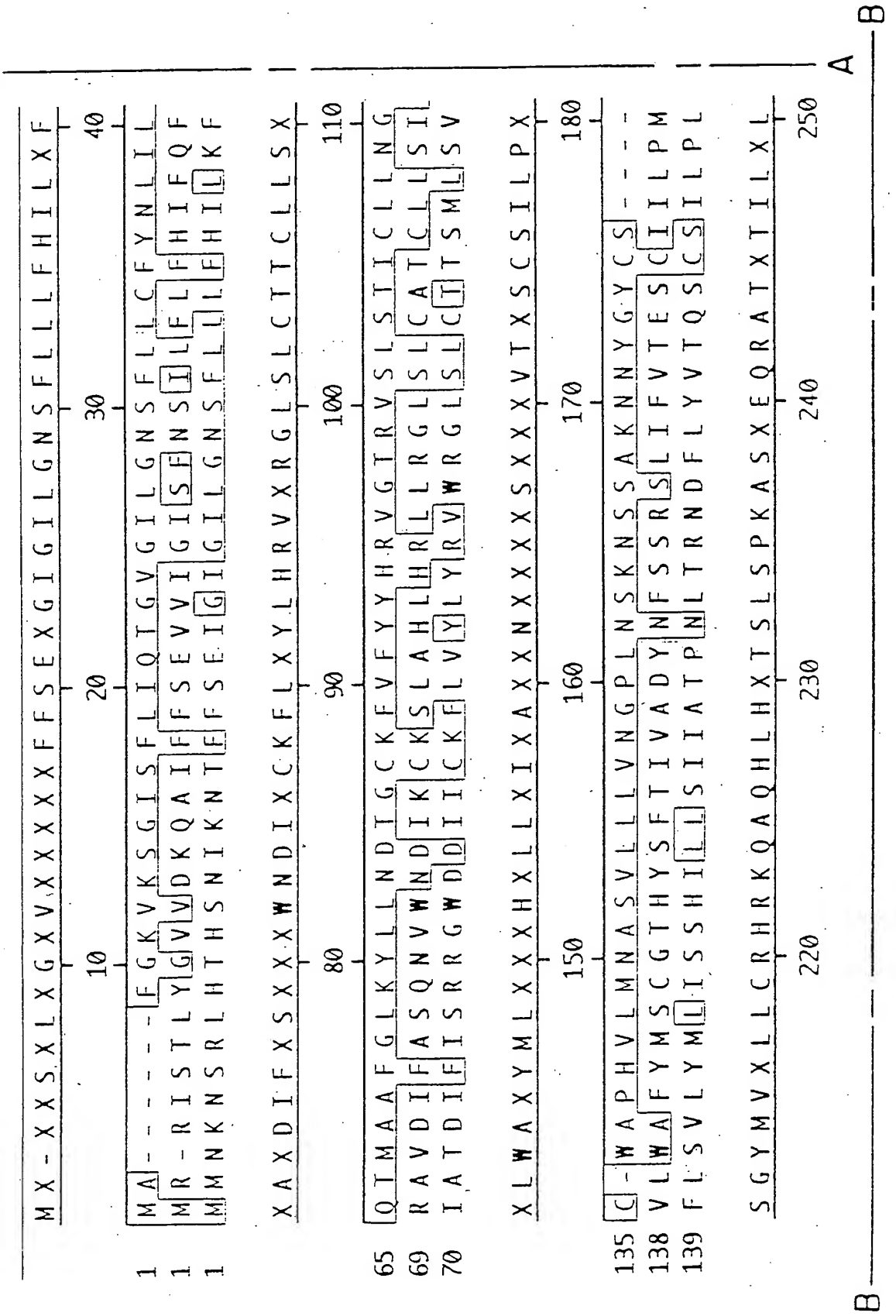
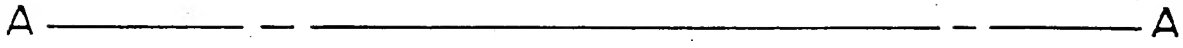


FIG. 2b



661/221 691/231
gca gtc tta tat ttt tcc cct gat ttt atg agt ttg ggc ttc atg gtc tgg gcc agt ggc
A V L Y F S P D F M S L G F M V W A S G
721/241 751/251
tcc atg gtc ttc ttc ctc tac aga cac aag cag caa gtc caa cac aat cac agc aac aga
S M V F F L Y R H K Q Q V Q H N H S N R
781/261 811/271
ctc tcc tgc aga cct tcc cag gaa gcc aga gcc aca cac acc atc atg gtc ctg gtg agc
L S C R P S Q E A R A T H T I M V L V S
841/281 871/291
tcc ttt ttt gtt ttc tat tca gtc cat agt ttt ctg aca att tgg aca act gta gtt gca
S F F V F Y S V H S F L T I W T T V V A
901/301 931/311
aac cca ggc cag tgg ata gtg acc aac tct gtg ttg gtc gcc tca tgt ttc cca gca cgc
N P G Q W I V T N S V L V A S C F P A R
961/321 991/331
agc cct ttt gtc ctc atc atg agt gat act cat atc tct cag ttc tgt ttt gcc tgc agg
S P F V L I M S D T H I S Q F C F A C R
1021/341 1051/351
aca agg aaa aca ctc ttt cct aat ctg gtt gtc atg cca tga
T R K T L F P N L V V M P *

FIG. 3c

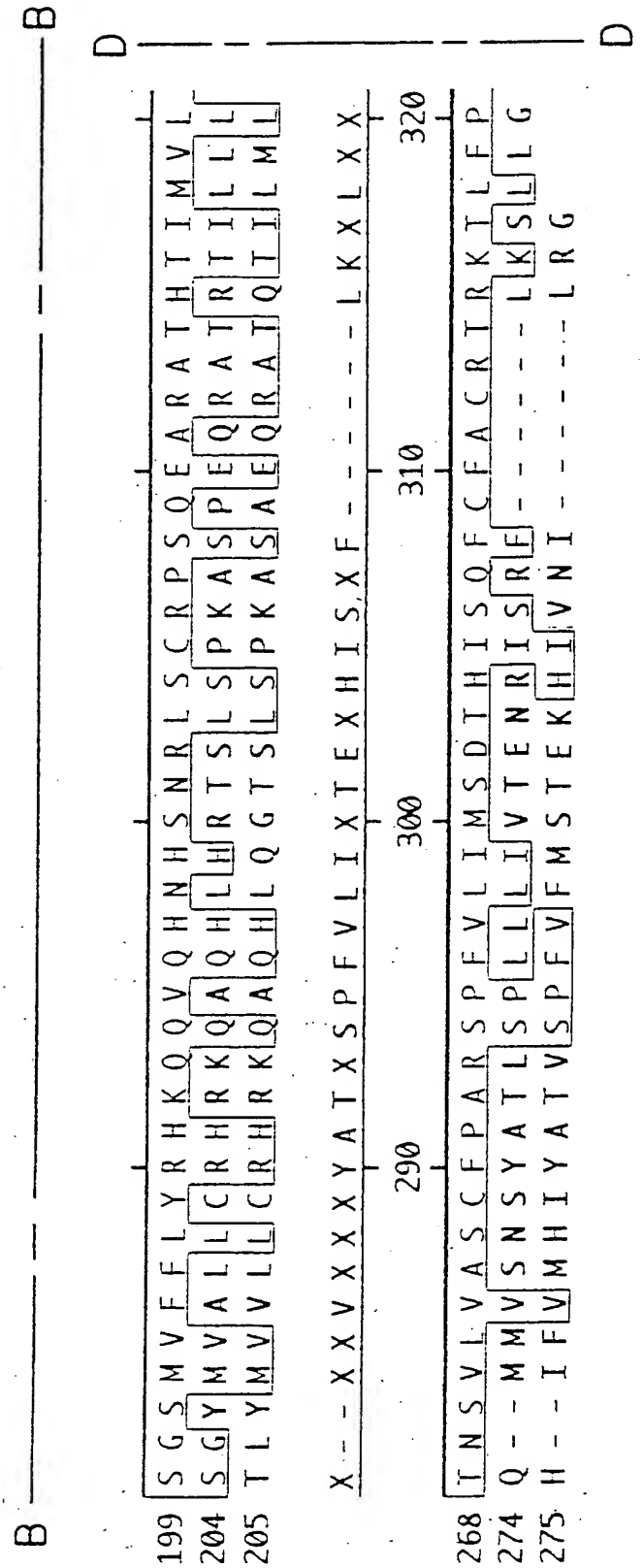


FIG. 3b

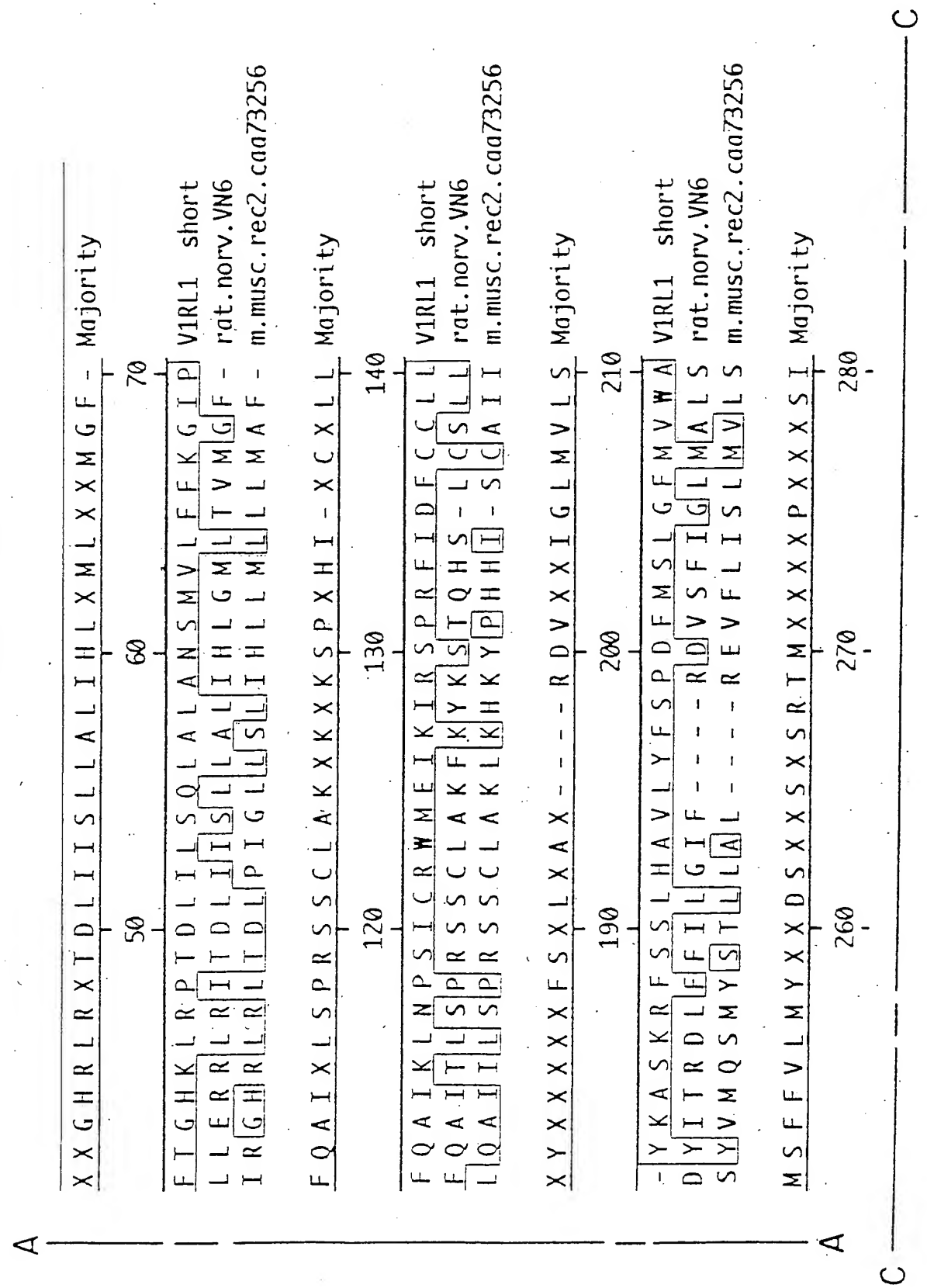

A
C
Majority
V1RL1 short
rat.norv.VN6
m.musc.rec2.caa73256
50
60
70
120
130
140
190
200
210
260
270
280

FIG. 4

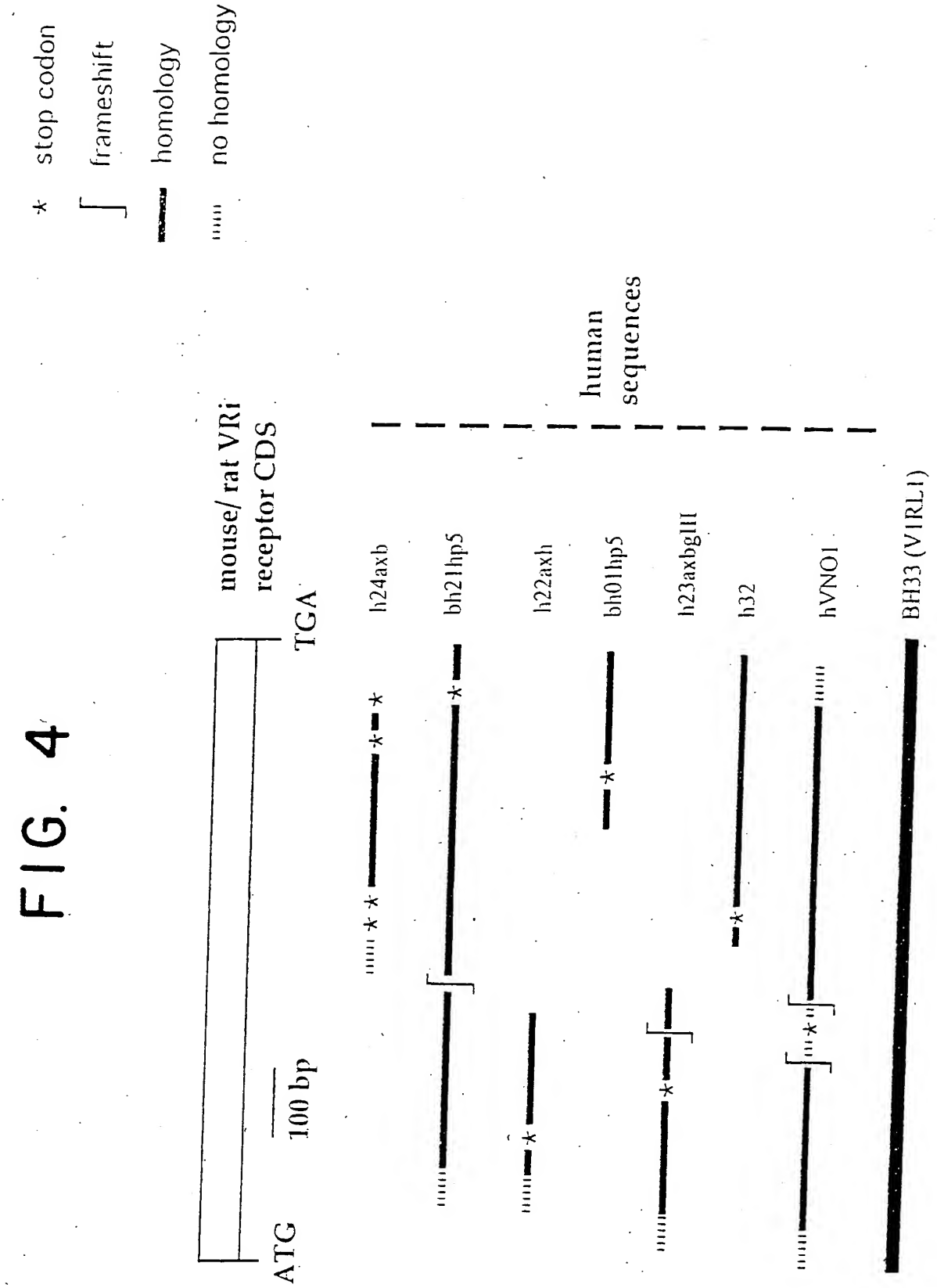
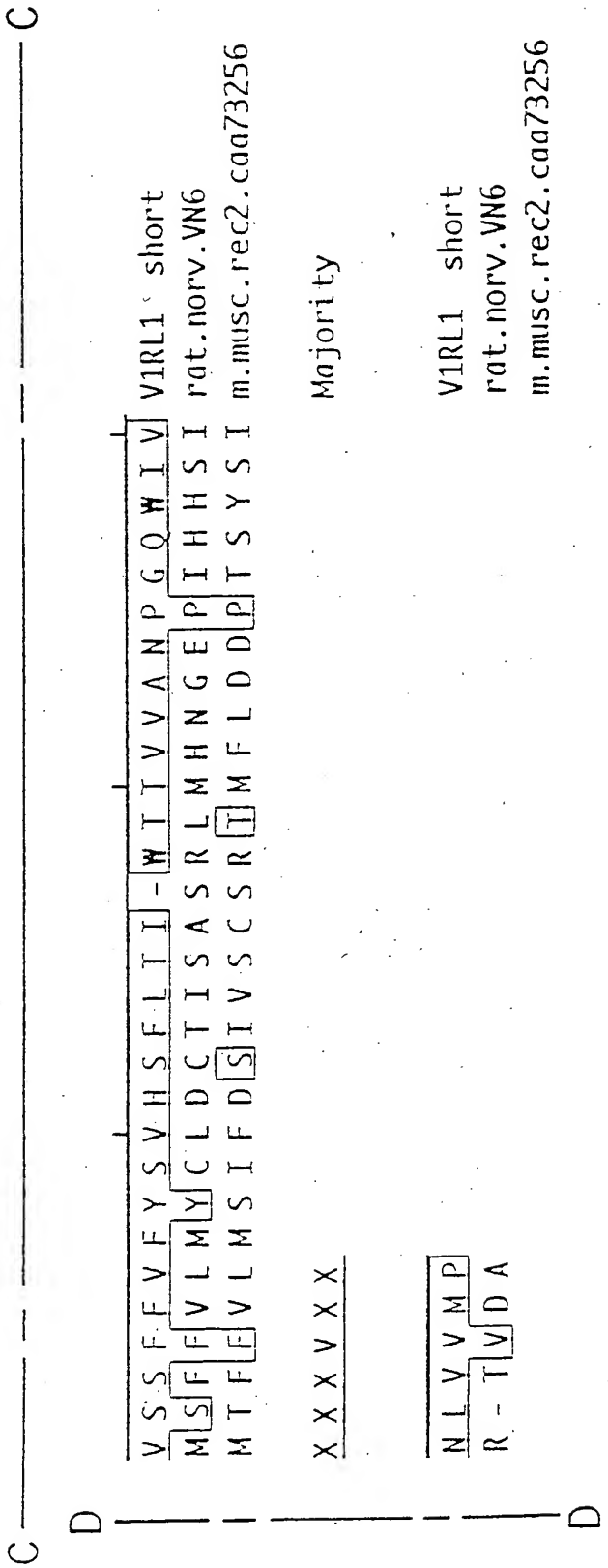


FIG. 3d


C
C
D
D
V S S F F V F Y S V H S F L T I - H T T V V A N P G Q H I V V1RL1 short
M S F F V L M Y C L D C T I S A S R L M H N G E P I H H S I rat.norv.VN6
M T F E V L M S I F D S I V S C S R T M F L D D P T S Y S I m.musc.rec2.caa73256
X X X V X X Majority
N L V V M P V1RL1 short
R - T V D A rat.norv.VN6
m.musc.rec2.caa73256

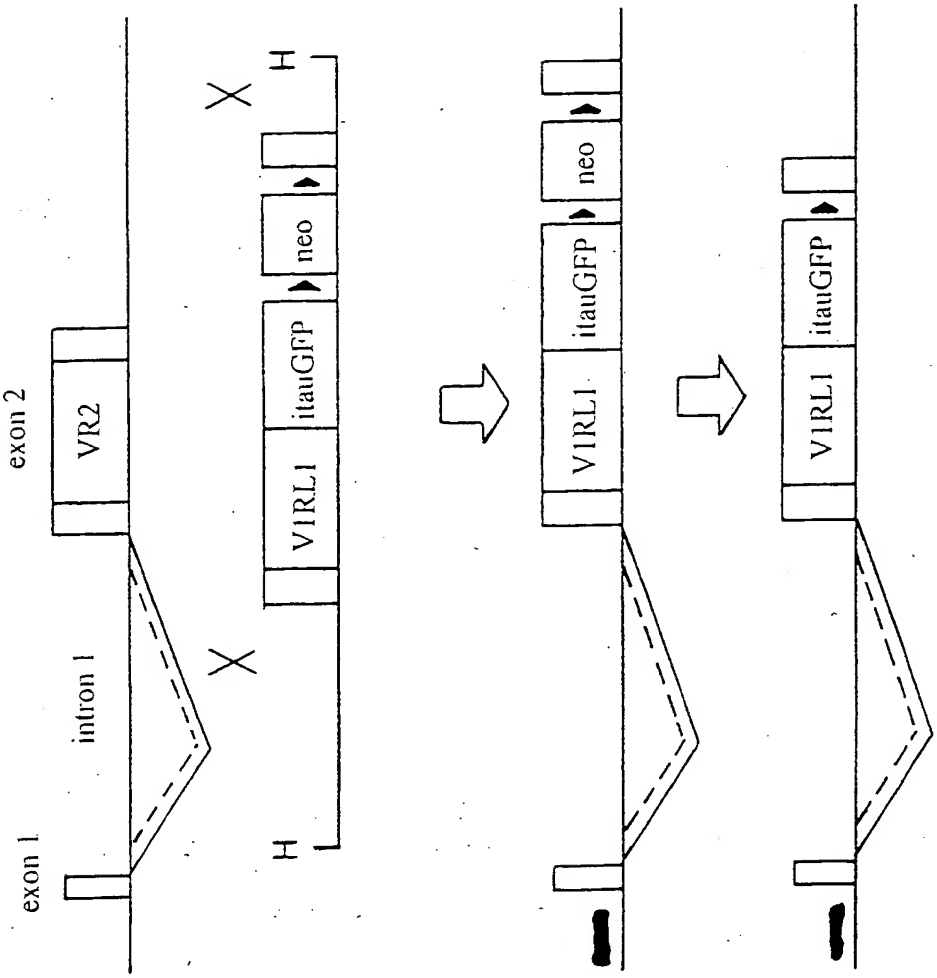


FIG. 6a

FIG. 6b

FIG. 6c

FIG. 6d

FIG. 5b

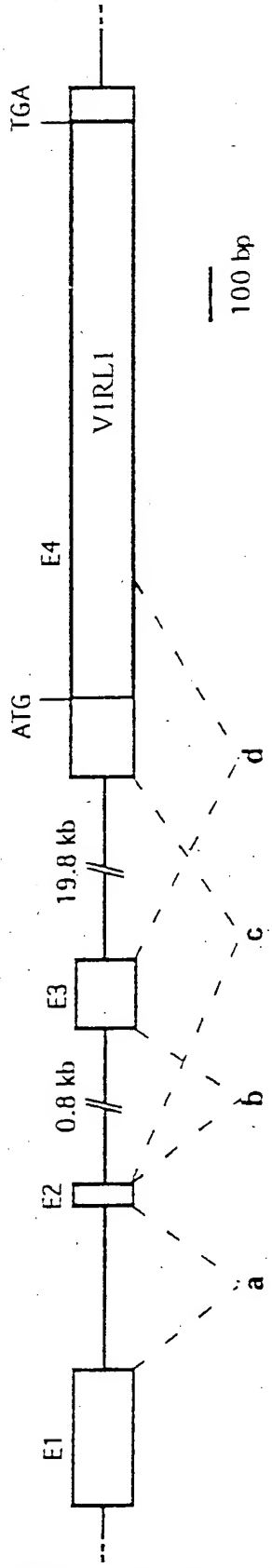


FIG. 7b

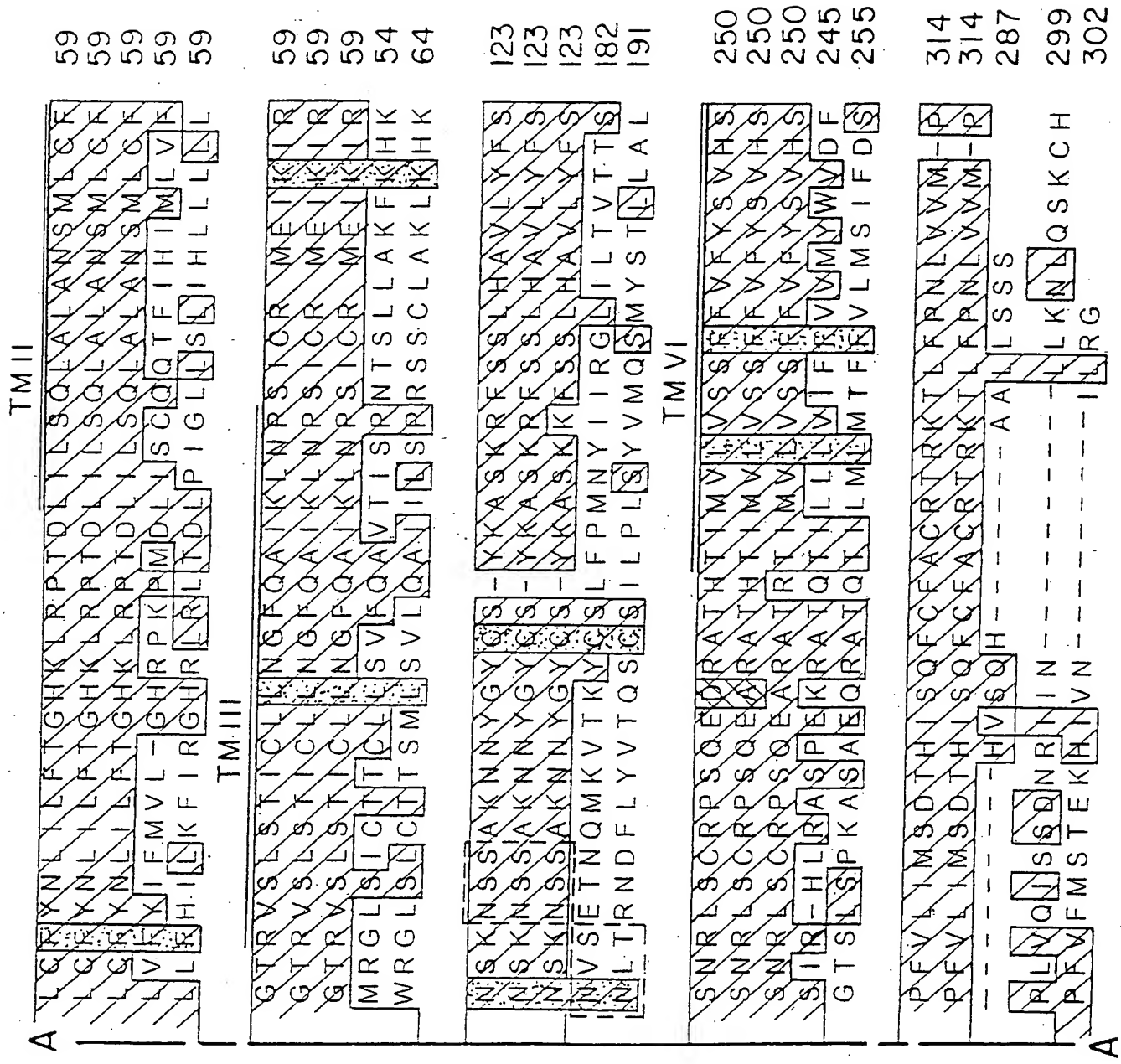
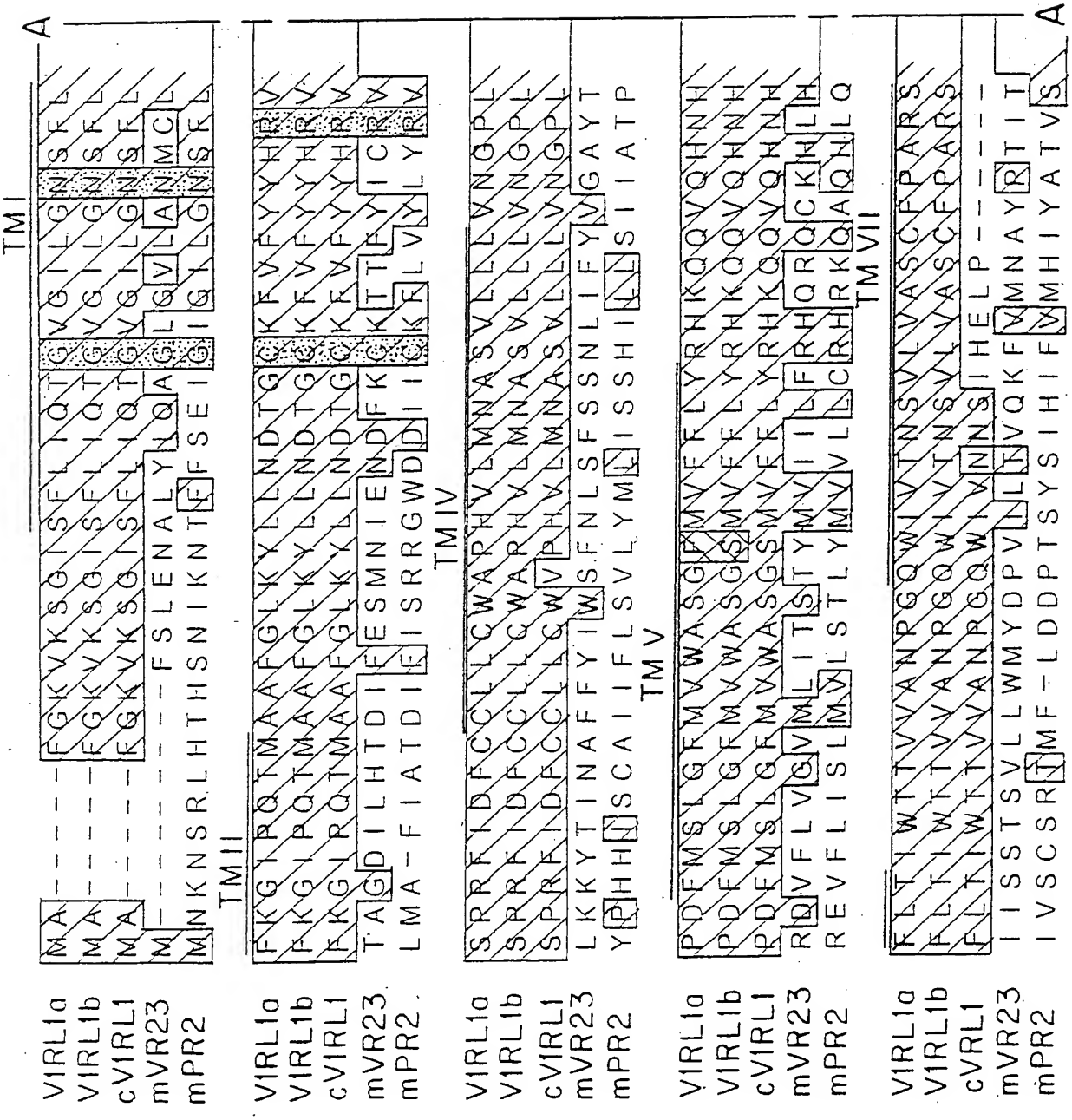


FIG. 7a



INTERNATIONAL SEARCH REPORT

Inter il Application No
PCT/US 00/27211

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/705 C12N5/10 C12N15/62 C07K16/28
C12Q1/68 G01N33/50

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, PAJ, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL SEQUENCE LIBRARY 'Online! 3 February 1998 (1998-02-03) LAMERDIN J.E. ET AL.: "Sequence analysis of a 500 kb ZNF gene family- containing human contig in 19q13.4 - unpublished" XP002157788 cited in the application accession no. AC004076 ---	10-15, 18,40-43
X	WO 97 14790 A (UNIV COLUMBIA) 24 April 1997 (1997-04-24) page 15, line 3 -page 16, line 2; page 19, line 12-14; page 48, line 27 - page 49, line 26; claims; SEQID15, clone HG25 --- -/--	1-5, 10-15, 17-31, 36-43

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

'A' document defining the general state of the art which is not considered to be of particular relevance

'E' earlier document but published on or after the international filing date

'L' document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

'O' document referring to an oral disclosure, use, exhibition or other means

'P' document published prior to the international filing date but later than the priority date claimed

'T' later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

'X' document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

'Y' document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

'&' document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
18 January 2001	06/02/2001

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Holtorf, S

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Inter il Application No
PCT/US 00/27211

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 99 00422 A (HARVARD COLLEGE) 7 January 1999 (1999-01-07) page 3, line 13; page 11, lines 11-19; page 15, line 27-32; page 16, line 1-4; page 18, line 13; examples, claim 11; SEQIDs 51-55	1-5, 10-15, 17-31, 36-43
A	WO 92 17585 A (UNIV COLUMBIA) 15 October 1992 (1992-10-15) page 32 -page 33	
A	DATABASE SWISS PROT DATABASE 'Online! 1 May 1999 (1999-05-01) SAITO, H.: "UNTITLED" XP002157789 ACCESSION NO. Q9Z196	
A	DULAC CATHERINE ET AL: "A novel family of genes encoding putative pheromone receptors in mammals." CELL, vol. 83, no. 2, 1995, pages 195-206, XP002157780 ISSN: 0092-8674 cited in the application the whole document	
A	MATSUNAMI H ET AL: "A MULTIGENE FAMILY ENCODING A DIVERSE ARRAY OF PUTATIVE PHEROMONE RECEPTORS IN MAMMALS" CELL,US,CELL PRESS, CAMBRIDGE, NA, vol. 90, 22 August 1997 (1997-08-22), pages 775-784, XP002913898 ISSN: 0092-8674 the whole document	
A	MONTI-BLOCH L ET AL: "THE HUMAN VOMERONASAL SYSTEM" ANNALS OF THE NEW YORK ACADEMY OF SCIENCES,US,NEW YORK ACADEMY OF SCIENCES, NEW YORK, NY, vol. 855, 1998, pages 373-389, XP000863095 ISSN: 0077-8923 figure 2	
P,X	RODRIGUEZ, I., ET AL. : "a putative pheromone receptor gene expressed in human olfactory mucosa" NATURE GENETICS, vol. 26, September 2000 (2000-09), pages 18-19, XP002157781 the whole document	1-5, 10-15, 18,32, 34-43

-/---

1

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
T	DEL PUNTA, K., ET AL. : "sequence diversity and genomic organization of vomeronasal receptor genes in the mouse" GENOME RESEARCH, vol. 10, December 2000 (2000-12), pages 1958-1967, XP002157782 the whole document	
T	KEVERNE ERIC B: "The vomeronasal organ." SCIENCE (WASHINGTON D C), vol. 286, no. 5440, 22 October 1999 (1999-10-22), pages 716-720, XP002157783 ISSN: 0036-8075 pges 716,719,720	

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

mation on patent family members

Inter ... al Application No
PCT/US 00/27211

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9714790 A	24-04-1997	AU 7450096 A	07-05-1997
		EP 0858505 A	19-08-1998
WO 9900422 A	07-01-1999	EP 0996635 A	03-05-2000
WO 9217585 A	15-10-1992	AU 669107 B	30-05-1996
		AU 1796192 A	02-11-1992
		CA 2106847 A	06-10-1992
		EP 0578784 A	19-01-1994
		JP 6509702 T	02-11-1994